STUDIES IN QUANTITATIVE INHERITANCE AND THEIR IMPLICATIONS FOR ANIMAL BREEDING.

by

Robert Caradog Roberts

B.Sc. (Wales)
Ph.D. (Edinburgh)

Thesis presented for degree of Doctor of Science
University of Edinburgh 1985

C O N T E N T S

SUMMARY

The thesis comprises a collection of 30 research papers, divided for convenience into four groups. They are studies in quantative genetics, with emphasis on the implications for animal breeding and the genetic control of production traits in domestic livestock. The experimental work, described in the first three groups, was all conducted on laboratory mice. The fourth group contains papers on non-experimental studies, where the application of the laboratory work to large animals is explored in more detail, with a few studies relating the same concepts to animal behaviour.

Group I, entitled "Genetic influences on reproduction", examines some aspects of the genetic control of litter size and its component traits, ovulation rate and embryonic mortality. The main study was on inbreeding depression and its converse, hybrid vigour. Their effects on the dam and on the offspring were examined separately. Inbreeding in the offspring had a greater effect in reducing litter size than inbreeding in the dam, but when heterozygosity was restored by crossing, crossbreeding in the dam increased litter size whereas crossbreeding in the offspring was of much less importance. The level of heterozygosity had little effect on ovulation rate, but inbred dams were less able to secure implantation. The material led to further examination of the mechanism of ovulation, and of some factors involved in embryonic mortality.

The papers in Group II report on the "Effects of selection on growth". The first five papers in the Group present the results of a long-term study of selection limits. An examination of the genetic nature of the limits showed that one line had no residual additive variance and had thus been driven to fixation, but this was not true of all lines. The most successful method of transcending the limit, out of several tried, was to cross lines previously selected in the same direction, though linkage impeded the response, Selection for the efficiency of food conversion, both at different ages and on different dietary regimes, led to the unexpected finding that the more efficient mice are fatter. Another paper reports that it is difficult to improve productivity by direct selection.

The next group of papers, Group III, is entitled "Genetic control of growth".

The papers all use lines selected for different weights and all attempt to answer the question : why do large mice grow more than small ones? It was found that the same body weight could be reached at different ages, showing that growth rate is not a simple function of mature size. It was also shown that small mice live longer than large ones. An extensive study of aggregation chimaeras between large and small mice yielded a disappointing result. The study failed to show whether there is a growth-controlling tissue or tissues. There could be, but if there is, its cell proportions are the same as in other tissues. A study of cell number and cell size showed that selection for large body weight increased both. The initial question had been whether cellular properties determined growth. The answer was that growth seemed to determine cellular properties; when organs were examined from large and small mice at the same weight, their cellular properties were virtually identical. But any similar suggestion of weight being the determinant did not apply to a study of food intake and efficiency in selected strains. Large mice both ate more and were more efficient than small mice, whether compared at the same age or at the same weight.

The last group, Group IV, is entitled "Non-experimental studies", carefully avoiding the suggestion of theoretical work, which it is not. Briefly, this section explores the relationship between laboratory animal research and animal breeding in practice. It also diverts into some of the implications for behavioural research, which may yet prove fruitful in addressing the concerns over animal welfare. Emphasis is placed on the use of laboratory animals to dissect complex phenotypes, in a way not readily available to those committed to multi-trait selection with domestic livestock. The section examines the progress still to be made in livestock with traditional methods, while it looks ahead to the routine adoption of gene transfer between species, though emphasising the biometrical requirements before such transfers can be exploited to good effect.

INDEX OF PAPERS

GROUP I: GENETIC INFLUENCES ON REPRODUCTION

1. R.C. ROBERTS. The effects on litter size of crossing lines of mice inbred without selection. Genet. Res. 1, 239-252. 1960.

2. D.S. FALCONER & R.C. ROBERTS. Effect of inbreeding on ovulation rate and foetal mortality in mice. Genet. Res. 1, 422-430. 1960.

3. D.S. FALCONER, R.G. EDWARDS, R.E. FOWLER & R.C. ROBERTS. Analysis of differences in the number of eggs shed by the two ovaries of mice during natural oestrus or after superovulation. J. Reprod. & Fertil. 2, 418-437. 1961.

4. J.C. BOWMAN & R.C. ROBERTS. Embryonic mortality in relation to ovulation rate in the house mouse. J. Exp. Biol. 35, 138-143. 1958.

GROUP II: EFFECTS OF SELECTION ON GROWTH

5. R.C. ROBERTS. The limits to artificial selection for body weight in the mouse. I. The limits attained in earlier experiments. Genet. Res. Camb. 8, 347-360. 1966.

6. R.C. ROBERTS. The limits to artificial selection for body weight in the mouse. II. The genetic nature of the limits. Genet. Res. Camb. 8, 361-375. 1966.

7. R.C. ROBERTS. The limits to artificial selection for body weight in the mouse. III. Selection from crosses between previously selected lines. Genet. Res. Camb. 9, 73-85. 1967.

8. R.C. ROBERTS. The limits to artificial selection for body weight in the mouse. IV. Sources of new genetic variance - irradiation and out-crossing. Genet. Res. Camb. 9, 87-98. 1967.

9. R.C. ROBERTS. Selection limits in the mouse and their relevance to animal breeding. Proc. 1st World Cong. on Genet. appl. to Anim. Prod., Vol. 1, pp. 493-509, Madrid, 1974.

10. W.K. AL-MURRANI & R.C. ROBERTS. Genetic variance in a line of mice selected to its limit for high body weight. Anim. Prod. 19, 273-289. 1974.

11. W.K. AL-MURRANI & R.C. ROBERTS. Maternal effects on body weight in mice selected for large and small size. Genet. Res. 32, 295-302. 1978.

12. E. YUSKEL, W.G. HILL & R.C. ROBERTS. Selection for efficiency of feed utilisation in growing mice. Theoret. & Applied Genet. 59, 129-137. 1981.

13. D.E. STEANE & R.C. ROBERTS. Selection for total weaning weight in the mouse, and its implications for domestic livestock. Z. Tierzucht. Zuchtungsbiol. 99, 222-231. 1982.

GROUP III: GENETIC CONTROL OF GROWTH

14. R.C. ROBERTS. The lifetime growth and reproduction of selected strains of mice. Hered. 16, 369-381. 1961.

15. R.C. ROBERTS, D.S. FALCONER, PATRICIA BOWMAN & I.K. GAULD. Growth regulation in chimaeras between large and small mice. Nature, Vol. 260, No. 5548, 244-245. 1976.

16. D.S. FALCONER, I.K. GAULD & R.C. ROBERTS. Growth control in chimaeras. Symp. on Genetic Mosaics and Chimaeras in Mammals. Gatlinburg, Tenn. pp. 39-49. Ed. L.B. Russell. Plenum Press. 1978.

17. D.S. FALCONER, I.K. GAULD, R.C. ROBERTS & D.A. WILLIAMS. The control of body size in mouse chimaeras. Genet. Res. 38, 25-46. 1981.

18. D.S. FALCONER, I.K. GAULD & R.C. ROBERTS. Cell numbers and cell sizes in organs of mice selected for large and small body size. Genet. Res. 31, 287-301. 1978.

19. E.J. EISEN & R.C. ROBERTS. Postnatal maternal effects on growth and fat deposition in mice selected for large and small size. J. Anim. Sci. 53, 952-965. 1981.

20. R.C. ROBERTS. The growth of mice selected for large and small size in relation to food intake and the efficiency of conversion. Genet. Res. 38, 9-24. 1981.

21. R.C. ROBERTS & NANCY R. MENDELL. A case of polydactyly with multiple thresholds in the mouse. Proc. R. Soc. Lond. B. 191, 427-444. 1975.

22. R.C. ROBERTS. Small eyes - a new dominant eye mutant in the mouse. Genet. Res. Camb. 9, 121-122. 1967.

GROUP IV: NON-EXPERIMENTAL STUDIES

23. R.C. ROBERTS. Some contributions of the laboratory mouse to animal breeding research. Part I: Growth. Animal Breeding Abstracts, 33 (3), 339-353. 1965.

24. R.C. ROBERTS. Some contributions of the laboratory mouse to animal breeding research. Part II: Fertility. Animal Breeding Abstracts, 33 (4), 515-526. 1965.

25. R.C. ROBERTS. Some Concepts and Methods in Quantitative Genetics. Chapter 11, pp. 214-257. In Behaviour-Genetic Analysis. ed. Jerry Hirsch, McGraw Hill, 1967.

26. R.C. ROBERTS. Some evolutionary implications of behaviour. Canad. J. Genet. Cytol. 11, 419-435. 1967.

27. R.C. ROBERTS. Side effects of selection for growth in laboratory animals. Livestock Prod. Sci. 6, 93-104. 1979.

28. R.C. ROBERTS. Genetic influences on growth and fertility. Symp. on Biology of the House Mouse, the Zoological Society of London, Dec. 1979. Symp. zool. Soc. Lond. No. 47, 231-254. 1981.

29. R.C. ROBERTS. Lessons to be drawn from selection experiments with laboratory animals. Proc. World Congress on Sheep and Beef Cattle Breeding, Palmerston North and Christchurch, New Zealand, 1980. Vol. 1, 253-259. Eds. R.A. Barton and W.C. Smith. Palmerston North, NZ: Dunmore Press. 1982.

30. R.C. ROBERTS & C. SMITH. Genes with large effects - theoretical aspects in livestock breeding. Proc. 2nd World Cong. appl. to Anim. Prod., Vol. 6, 420-438. Madrid. 1982.

REVIEW OF PAPERS

The majority of the papers in this collection refer to experiments on quantitative genetics with the laboratory mouse, and their application to problems of animal breeding. The application to domestic livestock depends on the relevance of the laboratory experiments to farm conditions. The first requirement is to choose traits that are important in farm animals, on the assumption that their main genetic features find parallels in the mouse. Given that, there are advantages in time and cost if the work is done on laboratory mammals. However, this limits the objectives that can be explored in laboratory experiments. If the objective were, say, to reduce backfat thickness in the pig, then there would be no point in doing the work on any animal other than the pig. But if the objective were to study the genetics of fat deposition, or its distribution between different depots, then we should assume that the way genes affect the physiology of both pigs and mice would be sufficiently similar to allow some general conclusions for the pig to be drawn from a mouse experiment. Closer still to primary gene action, biochemistry is biochemistry, or so the biochemistry textbooks would lead us to suppose. At the molecular level, the similarities between species are such that interchanges between them will soon become routine procedures, as they are already in lower organisms. Thus, the laboratory mouse, as a homeotherm that gestates and suckles its young, that digests its food and partitions the metabolites to various destinations, is a model for domestic animals. It is, however, only a model. Just how good a model, and with what reservations, is discussed in a group of papers at the end of this thesis, after the experiments on which the generalities are based have been reviewed. These experiments have been grouped, for convenience, into three categories. The first concerns various aspects of reproduction. The second examines the effects of selection on various measures of growth. The third, reporting mostly further experiments on the selected material, examines the genetic aspects of various mechanisms that may influence the control of growth. The final section, as suggested above, is an attempt at various syntheses of my own and other findings, viewed from various perspectives.

The papers are not presented in strict chronological sequence, but represent my attempt to place them in a more logical sequence to bring out the story. As explained earlier, the main motivation was the application to animal breeding, but this is not always stated overtly. The experiments were conducted more in the context of quantitative genetics, which is the parent discipline of animal breeding theory. A few of the papers, especially in the last section, concern the application of quantitative genetics to other fields of investigation. These particular papers involve an examination of the premises of quantitative genetics, which hopefully is of value to any application, including that to animal breeding. For the same reason, not all of the papers refer strictly to genetical phenomena, especially in the first and third sections. The line where genetics stops and something else begins is at best ill-defined. The point is that in applied work the line does not exist. If a problem in one conventional area leads to a solution in what may be a different area, we need not be too concerned about artificial boundaries. To the extent that genetics is now the core discipline of biology, we should perhaps expect that genetic aspects of litter size may lead to mechanisms of ovulation, or that selection for growth rate may end up in a discussion of voluntary food intake. Genetics is a tool as well as a discipline in its own right. Applications concern both the development of tools and their proper use once they have been developed, and this thesis does not attempt to distinguish between the development and the use. Both are parts of a more general scientific endeavour.

Group I. Genetic influences on reproduction.

This is a group of four papers deriving from the effects of inbreeding on litter size. The variation in litter size generated by inbreeding led to the examination of ovulation rate and embryonic mortality as separate aspects of the measured trait, and involve some examination of how these separate aspects are integrated.

Paper 1 is a study of the effects of inbreeding and crossing on litter size. At the time, although both inbreeding depression and heterosis

were well-known for litter size, there had been no definitive study of their complementarity. The study showed that in the absence of selection, the mean of the crossbred population was identical to that of the original outbred population. By now, this seems to be rather a trivial finding, and even at the time, the result could be deduced by logic. It did, however, furnish the first experimental proof that heterosis was not attributable to natural selection operating within lines during inbreeding. Later, Falconer was to show the dramatic effect of artificial selection between lines. Perhaps the main contribution of the study in Paper 1 was that heterozygosity in the dam and in the litter, separately, had unequal effects when inbreeding starts and when heterozygosity is restored by crossing. Initially, inbreeding in the litter had a much greater effect in reducing litter size than inbreeding in the dam. Conversely, when lines were crossed, crossbreeding in the dam had a much greater effect than crossbreeding in the litter. Thus, when the effects of heterozygosity are considered separately for the dam and its litter, those effects are markedly nonlinear. To this day, that nonlinearity, though amply confirmed since then, has not been adequately explained. The explanation given in the paper (1960) was that a maternal effect restricted litter size in inbred mothers irrespective of the heterozygosity of the young. That explanation in fact has a tautological element, but it has never been superseded by a better one.

In Paper 2, the effects of inbreeding and crossing are examined in more detail. By dissection of females in late pregnancy, ovulation rate can be estimated from the number of corpora lutea, while embryonic mortality can also be estimated, both pre- and post-implantationally. The results were very clear. Changes in heterozygosity had no effect on ovulation rate, which was the same for all levels of inbreeding. Somewhat fortuitously, the paper was able to show that ovulation rate was directly related to body weight, and not obviously to anything else. Inbred mice had been born and reared in smaller litters, and were consequently of the same weight as crossbred mice reared in larger litters. It was also the case that

an embryo once implanted, whatever the level of heterozygosity, had an equal chance of surviving to term. This, however, was not true of preimplantational mortality. Inbred dams were markedly less able to secure implantation whatever the heterozygosity of the embryo. It was speculated in the paper that this was due to some impairment of endocrine function in the dam, but proof was lacking. In retrospect, it is a pity that dissections were not done during the early stages of inbreeding, because we do not have a description of the effects of inbreeding on the embryo at that stage. Much later, I was reminded of the importance of synchrony between the gestational stage of the dam and the developmental stage of the embryo. That might have been a fruitful area for inquiry in inbred material, and for that matter still is.

Because the corpora lutea were recorded separately for the two ovaries, Paper 3, incorporating also some additional material, examines the control of ovulation. The empirical observation was that the correlation between the numbers of eggs shed by the two ovaries was negative after natural ovulation, whereas it was positive after the higher rates achieved by induced ovulation. This apparent contradiction was resolved in terms of the relationship between the mean and the variance of the total number of eggs shed. The switch in sign of the correlation was, in effect, a statistical artefact, and a little statistical development in the paper yields an elegant formulation of the correlation in terms of the mean and variance. The conclusion was that the ovary from which an egg was shed was essentially a random event, and that the overall distribution between sides conformed closely (though perhaps not exactly) to the binomial distribution. In other words, the system behaved as if (and I emphasise that it is only an analogy) a given amount of a circulating hormone was used up every time an egg was shed. This means that the total number of eggs shed behaved as if it was determined by the total amount of this hormone, and that the distribution of the circulating hormone between ovaries was left to chance. Despite a forbidding volume of literature on ovulation since that date, our conclusion, as far as

I know, has never been seriously challenged.

Paper 4 is another that stems from the basic observations in Paper 1, only this time it concerns embryonic mortality. It starts off as before with the observation of a negative correlation, after natural mating, between the number of eggs on the two sides. However, the negative correlation is not found in late pregnancy in terms of live embryos on the two sides. This points clearly to differential loss of embryos, the loss being proportionately greater as the number of eggs shed increases. The paper simply gives experimental proof of this, and further ascribes the differential loss to the preimplantational stage. This augments a finding reported in Paper 2: inbred dams with a preimplantational problem anyway would find it exacerbated by any increase in ovulation rate.

The four papers comprising this section are by now a bit dated, both in terms of the problems addressed and of the methodology adopted. Nevertheless, in reviewing them briefly, it is somewhat gratifying to find that their main conclusions still stand, and they are possibly the best account yet of the effects of inbreeding in a polytocous mammal.

Group II. Effects of selection on growth.

The pioneering work of Goodale and of MacArthur had demonstrated quite clearly the importance of genetic influences on body weight long before I entered the field. Following this early work, Falconer had refined the experimental techniques and derived good estimates of the relevant genetic parameters. Falconer's experiments were the first to show, at least for a mammal, that the response to selection did not continue indefinitely, though MacArthur had noted a tailing off of some of his responses. Falconer and King were the first to address the problem of what to do when the response to selection stops. They argued perceptively that because selected lines differ in some of their physical characteristics when they reach the limit, they may differ also in their genetic characteristics. Crossing such lines therefore ought to generate new genetic variance and yield further responses to selection. Their experiment proved their point. Later, these considerations were formalised in various treatments of fixation

and drift, but it may be fair to claim that Falconer and King were the first to demonstrate them experimentally. Reference to all this early work, if they are needed, will be found in the bibliography of Paper 5. The references to what follows are also there.

Work on selection limits was given new impetus by Alan Robertson's (1960) paper on "A theory of limits in artificial selection." For the first time, this paper predicted what to expect of a selection response, given certain assumptions, and replaced the empiricism that had characterised the interpretation of selection experiments until then. This is not to say that the empiricism had served the science badly. It had yielded quite reasonable estimates of how long a response would last and what magnitude of response to expect, as Falconer had summarised in the 1960 edition of his book. Robertson's treatment not only quantified the expectations but brought some new concepts to bear on selection theory. The first was an extension of Kimura's well-known formula for chance fixation. Thus, an allele may be fixed in a population by chance even though selection is directed against it. Robertson showed that this chance, when selection is based on individual measurements, is a function only of the product Ni, where N is the effective population size and i is the intensity of selection. The greater Ni, the less the chance of fixing unfavourable alleles. The second concept that Robertson introduced made it possible to test his theory against experiment. This was the introduction of the idea of the "half-life" of a response. It is impossible to determine when an asymptotic curve reaches its limit, but ex post facto, it is quite possible to read off the point in time when half the gain had been obtained. Robertson predicted that the half-life of a response would be obtained in 1.4N generations for genes that act additively, though this could rise to 2N for rare recessives.

This background has been given in a little detail (though not nearly enough to do justice to Robertson's theory) because it governed the

series of experiments reported in the next five papers. The time was obviously ripe for a new examination of selection limits, though the experimental programme I undertook was long-term and, frankly, it eventually proved to be quite tedious. However, if I am allowed one statement out of sheer vanity, it was amply rewarded when Gordon Dickerson, in a letter asking for the reprints, described the papers as "a definitive series where no more need be said". Though in retrospect this was clearly an over-generous tribute, I have kept that letter and shall always treasure it.

The first paper in a series of four is listed in this collection as Paper 5. This presents re-analyses of previous selection experiments for body weight in the mouse, with the overt intention of testing Robertson's theory. In the event, the theory stood up reasonably well. The parsimonious conclusion of the paper was that the exhaustion of the additive variance was a sufficient explanation of the limits attained. By now, I am not certain that I should care to defend this conclusion too strongly, and to the extent that the conclusion may not be entirely justified, Robertson's theory can accommodate some margin of error without being disproved. Nevertheless, the theory helped substantially in the interpretation of the results. The very short half-lives indicated that most of the favourable alleles had been fixed. This helped to explain why four large lines had reached rather similar limits, as had three small lines. Some messy algebra had also provided estimates of gene numbers and of proportionate gene effects, which (to use a double negative quite deliberately) were not unreasonable. However, if I were to repeat the exercise, I should probably manipulate the data to yield higher estimates of gene numbers and reduce the proportionate effects accordingly. The system is sufficiently flexible to allow this kind of manipulation without serious challenge.

One conclusion from this study now strikes me as being rather important, and though stated in the paper, it was rather lost among the other conclusions. This was that for populations of that size, the

selection intensity was close to the optimum for securing the maximum gain. As this had been combined with rapid initial responses, the point has considerable relevance to the breeding of domestic livestock, and one which perhaps does not receive sufficient attention.

The genetic nature of the limits is explored in Paper 6, establishing that in one large line, the additive genetic variance had indeed been exhausted by selection. A small line, in contrast, responded readily to reversed selection, and the astonishly high realized heritability of the reversed response indicated that considerable additive genetic variance had remained in the line at its limit over a period of 20 generations, despite continued selection for small size. The retention of so much genetic variance in a line at its limit demanded explanation. It is argued in the paper, somewhat circuitously though it still reads quite convincingly, that the genetic variance was retained because of natural selection operating through mortality. Thus, a balance had been struck between the positive deviations in weight of the animals that were selected and the negative deviations of those that died. However, to make the argument tenable (because relaxed selection had yielded little response) it had to be postulated that a critical weight had been reached in the small line, above which natural selection did not operate. While this postulate may not defy credulity, it perhaps puts it under some strain. However, the main conclusions of the paper are quite clear. Its weakness is that it lacks generality, because only one line in each direction was analysed. This weakness was realized at the time, and the paper ends with a plea for more experimental work. But except for Falconer's work on the limits reached in a line selected for litter size, and a similar study on body weight reported in Paper 10 of this collection, I am not aware of any further work on the nature of the limits in the mouse.

The experiments described in Paper 7 were an extension of the approach used earlier by Falconer and King to break through the limit and secure further responses to selection. The methodology therefore was not new, and the principle had already been established. Nevertheless, the results were not entirely predictable and a new feature of

linkage was brought out. Four large lines were crossed to form a base population to select for larger body weight, and three small lines were used similarly to form a base for selection for low body weight. Heterosis in all first crosses showed that the lines selected in the same direction all differed at least in allelic frequencies. The heterosis was greater than Falconer and King had found, and more selected lines were put into the pool, both implying the potential for more genetic variance. It is therefore not surprising that renewed selection for larger body weight showed a substantially greater gain than Falconer and King had obtained. Less easy to explain was the poor response obtained in selecting for low body weight in similar circumstances. After 24 generations of further selection, the low weights of two of the original lines had not been recovered. The data from this and other experiments of the series did not allow for any general increase in weight for environmental reasons. The poor response of the low line is extensively discussed in the paper, and it reads as unconvincingly today as when it was written. Looking at the results again, I perhaps missed the most obvious explanation, or perhaps I could not bring myself to admit it. That explanation is that after 24 generations - shall we say, six years - I ran out of patience and terminated the experiment prematurely. The response for the last nine generations, in retrospect, looks reasonable good, or at least better than it had been. This of course only shifts the question to why the response had been so poor up to generation 15. In the paper, I seem to have argued myself out of acceptable reasons for the unimpressive total response. This must have satisfied the editor of the journal as well, for he did not bring me to task for my possible failure to take the line to its new limit!

The renewed selection from the large crosses had also shown a lag in response, but this time only for about 6 generations. This led to a discussion, and quite a reasonable experimental proof, of the influence of linkage in crossbred material. The paper reads as if I had been rather alert in setting up this proof, but in truth, it was probably more good luck in guessing what the outcome might be.

The result however led to what now seems to be a percipient discussion of theoretical considerations, and to my mind this is the best section of the paper. The discussion starts with the statement of Hill and Robertson (reference given in bibliography of Paper 7) that unfavourable alleles are more likely to be fixed if they are linked to favourable ones. The lag in response in my experiment is then explained in the following terms. Unlinked loci in the original lines would have been fixed for the more favourable alleles, as chance fixation was largely excluded (Paper 5). Furthermore, some overlap in origin increased the chance that these fixed favourable alleles could be the same ones in different lines. When the lines were crossed, such loci might thus not contribute much to the new genetic variance. Contrast these loci with the situation at linked loci, where Hill and Robertson suggested that unfavourable alleles could be fixed, by chance. When the lines were crossed, recombinants among linked loci would not only generate new genetic variance but would be its main source. The experimental results were perfectly compatible with this model, as it took a little time before recombination generated sufficient variance to yield a response. But from the beginning, enough variance was released for the artificial selection to counteract the natural selection that had occurred in an unselected sample of the same material. We thus have a satisfying explanation for an unexpected result, with the further conclusion that this amount of variance from linked loci is a peculiarity of crossbred material, and of no other.

The difficulty of applying the same solution to the even slower response of the low line was claimed in the paper to be two-fold. First, the linkage had to be much tighter, and this does not seem entirely plausible. Second, and adding to first, alleles for small size tend to be recessive, and selection for such alleles should be more effective. A third reason seems to have been ignored in the paper. This was that the small lines may not have been fixed anyway, as shown in Paper 6. This fact would surely have reduced any impeding effect of linkage.

It was thus shown that crossing previously selected lines was a practical method of breaking the limit for large size, if not for small. Paper 8 explores other methods of introducing new genetic variance. The first method to be tried was gonadal irradiation of male mice. There had been some earlier irradiation experiments with Drosophila, where the findings were at best variable and at worst contentious. My experiment with mice did nothing to resolve the situation. There may have been a little response to subsequent selection, perhaps, but this could not be established with any degree of confidence. With hindsight, it is easy to establish that the experiment was ill-conceived, because the response obtained if any was at least as good as might have been expected. Even moderate foresight might have established the same point. But it may have spared someone else from committing the same folly. However, the largely negative result led to a very clear conclusion. It was that a single dose of 600r. was not enough to give anything worthwhile in terms of increased genetic variance. At the very least, repeated doses should have been used. I also suggested that fractionated doses might be used to increase the mutation rate. Above all, there was the tantalizing hint that the result may not have been entirely negative. However, I may have spelled out the difficulties of this approach too clearly for any further attempt to break through the limits in mammals by using irradiation. To the best of my knowledge, this was never subsequently attempted, and the use of irradiation for this purpose seems to have gone out of fashion for any species. By now, experimenters have more faith in the powers of molecular manipulation, though as yet the procedures are far from routine. In reviewing my paper, the attempt now strikes me as pretty feeble, but it should have helped to design better ones. We must therefore suspend judgement on the amount of genetic variation in quantitative traits that could, in principle, be generated by irradiation.

The second method described in Paper 8 of generating new variance in a line at its limit had some success. This was to cross a line at its limit to an unselected line, or what in animal breeding terms might be called an unimproved line. The gains obtained by further selection from this cross, in two separate lines, were less impressive

than those described earlier from crosses between selected large lines. As a way of improving domestic livestock, crossing to unimproved stock, though perhaps feasible, carried a further penalty, as it took me about 9 generations to recover the original level of the selected line. But the experiment has two features of interest. First, a lag in the response in both lines confirms the phenomenon of linkage in such material, as discussed earlier. Second, the experiment amply demonstrated that there were alleles in the unselected line that were better than those in the selected line, and that the gain from their total effect was well worthwhile. This offers both hope and challenge to the application of molecular techniques to animal breeding. There are desirable alleles available. The problem is first to identify them, and then to incorporate them in the properly regulated pathways of the improved genotypes. At the time when I reported the experiment, this kind of guidance for the future did not awake my interest. If it had, I might have come out with more positive conclusions.

These four papers on the limits to selection have perhaps been given more space here than they deserve, but they represented a major experimental undertaking at the time and I find it instructive to reappraise them for any lasting relevance. The genetics could obviously have been done better, or at least more fully. They also illustrate how the same data can lead to modified conclusions as the science develops, and perhaps also as the passage of time erodes personal involvement and thereby lends more objectivity to the exercise. What does not seem to have changed is the potential application of the results to animal breeding, though that may be largely because no call has yet been made for that application in practice. The lessons for animal breeding, and the difficulties, are discussed in Paper 9. This paper might perhaps have been consigned to Group 4 (later), but it contains a summary of two important experimental findings. Looking back now, I cannot understand why each of these two findings was not the subject of a separate paper, because both are clearly of sufficient importance. Perhaps it was a case of pre-empting fuller publication by incorporating new material in a general review for an international congress. There is a temptation to find something original to say under such circumstances, but having done so, this clearly removes the incentive

for proper publication in a refereed journal.

The first important point to be found in Paper 9 affects the interpretation of virtually all the conclusions in Papers 6 to 8. It so happens that by good luck, the interpretation adopted earlier was correct, but this is no excuse for failing to document the facts fully when they become available. In Papers 6 to 8, the control line used was the largest of four lines that had earlier reached a limit. The interpretation of several issues was complicated by the fact that after the new experiments had begun, the line kept as a control had shown a tiresome rise in mean of about one phenotypic standard deviation. It was argued at the time that this increase was a genetic change, unique to that line, and should essentially be disregarded. It is shown in Paper 9 that the rise in mean was indeed a genetic change unique to that line, either a new mutation or a rare recombinational event. Those two possibilities could not be distinguished. It is also mentioned in Paper 9 that although a unitary genetic change must have been the explanation, it was a hopeless task to demonstrate segregation in such material. The story should have been told in full.

The second finding placed obscurely in Paper 9 is even more important. Lines selected for large body size are characterised by many sterile matings, because the females become too fat before mating. This experience is not limited to the laboratory mouse. Having lost two lines from this cause, I took an offshoot of one just prior to its extinction and mated it at a younger age, before fat accumulated. This offshoot kept going at a high level of fertility for a further 23 generations. This indicated clearly that the infertility of the original line was attributable to its excess fat. It was not a consequence of accumulated inbreeding nor the effect of increased body size as such, nor was it any other consequence of the selection programme that might have led to infertility as a correlated change. It is true that some of these separate effects had been established from earlier experiments, while the remainder required no more than intelligent guesswork. But as an integrated analysis of a practical problem, the solution should have been better reported than it was.

My final experimental attack on the limits to selection for body weight is reported in Paper 10. This was an attempt to apply a method successfully developed by Falconer (reference in bibliography of Paper 10) for improving litter size in a line at its limit for that trait. Briefly, Falconer had argued that recessive genes impeded further response to selection, as selection would become increasingly less successful in reducing the frequencies of such genes. The practical question was what to do about such a line if no similar line was available for crossing, thus removing the option described earlier in Paper 7. Falconer's answer was to derive a number of inbred lines from the selected line at its limit, to select between them and to cross the best of the survivors. The result was to improve litter size further by some 17 percent. Quite clearly, the method had to be tried for body weight. As Paper 10 shows, the result as far as improving body weight is concerned was a failure, but in a sense, the experiment validated the methodology. It showed that the premises were true, but that their exploitation is governed by the genetic properties of the trait in question. For litter size, Falconer had been able to postulate recessives at some 40 loci, each with an effect of about half of a phenotypic standard deviation and at a frequency of around 0.2. Recasting the numbers game in Paper 10 to genes of the same effect and at the same frequency, the number of loci still segregating in the body weight line was only about a third of that found by Falconer for litter size. It was further calculated that even if all of these recessives for body weight had been eliminated, the total improvement would have been a negligible 2%, which no reasonable experiment could hope to establish. Two points emerged from this. The first was that in a line at its limit, single gene effects of substantial magnitudes could remain without detectable effects on the genetic variance. The second point was that a method successfully used for litter size could not be applied to the more additive trait of body weight. These two points together lent generality to Falconer's method, and to that extent, the experiment served a useful purpose, though its outcome at the time was something of a disappointment.

The experiment also qualified a conclusion derived in Paper 6, where a large line at its limit was shown to be fixed. This experiment was

on another large line, also at its limit, but which had not up to that time been driven to fixation, as was clearly shown by differentiation among the derived inbreds. This emphasises the danger, if any emphasis is necessary, of drawing general conclusions from a single unreplicated selection experiment. The work reported in Paper 10 would have gained if more use had been made of the replicated selected material by then available. That material was not exploited because of lack of space and labour.

The next paper could be taken more logically in conjunction with Paper 19 (later), but will be reviewed briefly here because it follows naturally from Paper 10. Paper 11 addresses the question of maternal effects in selection responses for body weight, and at least by implication, the role of maternal effects in determining the limits. The well-established technique of egg transfer was used. It is instructive to re-examine the data with specific reference to the influence of maternal effects on the limits, because the emphasis in the paper seems to have been somewhat different. Prenatal effects were judged to be only of minor importance, except for those operating through litter size. Even so, the uteri of small mothers seem to have been slightly inadequate for large foetuses. Small foetuses had equal birthweights irrespective of the genotype of the host mother. After birth, the corresponding effect was much more pronounced. Large offspring showed substantially reduced weights when reared by small dams. The superiority of large dams, though still detectable, is much less marked for small offspring. The conclusion is clear. Even though the selection for body weight had been done within litters, calculated to remove some of the complications of maternal effects, such effects were still detectable with respect to uterine and, especially, lactational performance. These aspects may of course be nothing more than a direct consequence of body size, and as such, it may be somewhat misleading to regard them as maternal effects. However, the formal distinction between the consequences of body size and of maternal effects *sensu stricto* cannot be drawn in this material. Whatever interpretation we choose, the conclusion in the paper is that the contribution of these genetic maternal effects, so-called, to the total response (at the limit) "was at most 20% and generally..... somewhat less". I cannot quite reconstruct any more the thinking behind this

rather dismissive statement; 20% now strikes me as a considerable amount. In any event, whereas the paper clearly showed small mothers to be maternally inadequate for large offspring, this is not the same as finding large mothers to be fully adequate. In fact, there is a hint to the contrary, in that large mothers coped better with five offspring of their own strain than they did with ten, with some residual effect on the weights of the offspring even at six weeks of age. It is not impossible therefore that selection for large size may have been somewhat impeded by maternal capacity not fully keeping pace with the increase in weight of offspring.

In fairness to my co-author, I should point out that in reviewing Paper 11, I have departed somewhat from the main issues considered at the time. The paper in fact describes a novel use of egg transfer and certainly brought out some clear differences in the maternal performances of large and small mice. That the data now allow us to reflect further on the nature of the limits is a tribute to their wider significance.

The two remaining selection studies to be reviewed in this section were both on topics of direct relevance to animal breeding. The first, described in Paper 12, was on selection for the efficiency of feed utilisation. The farm animal interest is reflected by the - presumably subconscious - use of "feed" for "food" in the title. It was a comprehensive study using four treatments: two age intervals and two feeding regimes. In addition, selection on each treatment was replicated. This sophisticated design was marred by the fact that the control lines, though set up and kept, were measured only sporadically. The occasions when measurements on the controls were made are identified rather lamely in the paper as "when spare capacity was available at the right time". The capacity should have been budgeted better, because the lack of contemporaneous controls rendered some of the data useless. Though little was made of it in the paper, a finding of major importance was the variation in mean efficiency over time. Some possible causes for this suggest themselves immediately, like a fall in ambient temperature which increases

food intake, or variation - sometimes all too obvious - in the physical quality of the diet. The lesson is obvious. Efficiency as a measurement is perhaps more susceptible than most to environmental influences, and as such, is more demanding than most of adequate controls. The paper would have done well to highlight this point, rather than dismiss a flaw in design with glib phraseology.

Flaw in design notwithstanding, the experiment showed quite clearly that it was possible to improve efficiency by selection, at both ages and both on ad libitum feeding and when a fixed amount of food was fed. This result, and most of the correlated changes, served mostly to consolidate and draw together similar results from other workers. The implications for domestic livestock are discussed, and situations are identified where direct selection for efficiency may be better than selection for gain alone, where improved efficiency generally follows as a correlated change. However, one correlated change was totally unexpected; mice selected for efficiency became fatter. This finding challenged the conventional bioenergetic view that it is more efficient to lay down lean than fat. That view could in any case be partly fallacious, as it is based on the combustible energy of the two tissues. The relevant energy here is the one for synthesis, and protein synthesis generates much heat which is lost. This and other technical aspects are discussed in the paper. In its bare essentials, the conclusion is that if heat dissipation and fat accretion are alternatives, it is more efficient to lay down fat than to lay down nothing at all.

Another unexpected finding was that when mice selected on one feeding regime were tested on the other, there were no interactions between feeding regime and efficiency. In other words, the genes promoting efficiency on ad libitum feeding were the same genes that promoted efficiency on a fixed intake. Quite why this should be so is somewhat mysterious, because on ad libitum feeding, genes controlling appetite could reasonably be expected to affect the outcome. The paper fails to explain why changes in appetite did not happen, but it does draw attention to the lack of any close connection between appetite and efficiency, which is another point of practical importance.

of the paper
The last sentence now strikes me as singularly poignant. Paraphrased, it suggests that several years and a few million pounds from now, a commissioned experiment by the Ministry of Agriculture, Fisheries and Food on systems of testing for pigs "might prove to be superfluous". Such, apparently, was our faith in the generality of our results. I rather wish now that we had used the same results to suggest the need for a similar experiment with pigs. It would not have been a better conclusion, but by now it would have read that much better.

The last experiment in this group of papers, reported in Paper 13, was an attempt to improve productivity, as defined by the total weight of the litter at weaning. Selection was not effective in improving the trait over 18 generations, and any apparent gain over the next two generations, before the experiment was terminated, should be discounted. Perhaps the experiment should have been continued, but that is simply being wise after the event. In practice, a selection programme that requires 20 generations to show any gain is of no interest anyway. The experiment at least did suggest why the selection programme adopted was not successful. This is of interest to breeders of domestic livestock, where productivity is a trait of major importance but, as in the mouse, generally intractable to selection on any direct measurement. The paper discusses reasons for this in some detail. In brief, the main difficulty seems to be that the trait as measured, the total weight of the litter at weaning, is an inappropriate index of its two components - number weaned and the mean weight of those weaned. These two component traits are negatively correlated environmentally, for obvious reasons, and due allowance should be made for this when selecting. Some theoretical work quoted in the paper, and which had been inspired in part by the result of this experiment, suggests that an appropriate index might be calculated from the number born and the subsequent growth rate of the litter when standardised to a given size. Despite the negative result from this and similar experiments in the mouse, the analysis in Paper 13 suggests that genetic improvement of productivity in domestic livestock may yet be obtained if the appropriate methodology is adopted.

Group III. Genetic control of growth.

The emphasis in this group of papers changes from direct genetical operations, where the objective was basically to change gene frequencies in a desired direction, towards attempting to understand how those changes in gene frequency brought about the desired effects on the phenotype. Obviously, there are several levels at which this kind of question can be asked. The question in this group of papers are all variants of the following: what is it about large mice that makes them grow more, and small mice less? Sometimes that question is asked directly with respect to some specific component of growth. At other times, the approach is the less direct one of examining other attributes of large and small mice, and attempting to relate these attributes to body size.

Paper 14 examines the lifetime growth and reproduction of large and small mice. The paper was the first to show that the same mature weight could be reached at different ages, though there had been no intention to change the shape of the growth curve. But the paper's main contribution was to show that small mice lived longer than large mice, and that this was particularly true of the length of reproductive life. The small lines had more litters (by a factor of 2 or 3) and weaned at least twice as many offspring as the large lines. Reading this paper again after a long time reminds me that genetic homeostasis was still in vogue then, and the Introduction to the paper is quite inappropriate to what follows. The stage is all set to expect lines selected in either direction to be less fit, and nothing of the kind happens. Having thus set myself a problem in the Introduction, I have to use some of the Discussion to argue myself out of it. I could have saved myself some time by not writing either. However, it must have made me think, because I wrote a very sensible suggestion of why small mice had a longer lifespan and particularly, a longer length of reproductive life. Drawing on the findings from some nutritional work, where a reduced calorific intake lengthened life and delayed the age of reproductive failure, I argued that small mice represented a biological means of restricting calorific intake, perhaps by diverting proportionately more of that intake to heat production. While we still lack proof of the effect on longevity, subsequent work on aspects of food intake and thermogenesis in small mice has strengthened the original suggestion.

The next three papers (15 to 17) are successive accounts of a long-term collaborative study, of which we had high initial hopes but which in the end proved to be somewhat disappointing. The basic approach was to make aggregation chimaeras between strains differing in body size, to investigate the influence of different cell lineages within an animal on its growth. The technique never worked smoothly in our hands, but thanks to the perseverance of my colleague, Dr I K Gould, the total number of chimaeras eventually produced makes it one of the larger chimaera studies on record.

The questions asked of the study by now seem to become progressively more sophisticated as the account of it develops, but as I recall them, our initial thoughts, naïve though they may have been, had the apparent virtue of a simplicity which could not fail to yield answers. We had imagined that there was, somewhere, a growth-controlling tissue or tissues which, in chimaeras, could derive from either the larger of smaller component strains. If the tissue derived from one or the other, body size in chimaeras would be distributed bimodally. If the tissue derived from both, body size would be intermediate. That approach has a kind of irrefutable logic about it that leads easily to the next step : by identifying the extent of the chimaerism in different tissues, the growth-controlling one (or more) could be identified from its correlation with growth. I do not recall that we ever did image that chimaeric mice would be lumpy, so from the start, a systemic control of growth was assumed. In fairness to my co-authors, I should perhaps not brand them with equal simple-mindedness, because this simple framework did in fact fail to produce good answers. We did not establish that there was a growth-controlling tissue, or if there was, we had no way of finding it. The distribution of chimaera body weights was neither biomodal nor centrally distributed, but something that more or less uniformly spanned the range. And we were left with the suspicion, at least, that chimaeric mice were more lumpy than they appeared. So much for simple questions. Perhaps the main value of the study was to show that chimaeras, by their nature, were not suitable for the purpose we had in mind. This realisation evolved gradually over the successive accounts of the programme.

The first paper of the three, Paper 15, was written in an obviously

optimistic mood, and reflects the simple yet clear thinking behind our first ideas. Everything was still on course, and the preliminary results gave some clear answers. The measurement of chimaerism was derived from the pigmentation of the patchy coat, as the component strains were of distinctive colours. It was a subjective measure, but it worked well. It showed that body size was linearly and directly proportional to the cellular composition of the pigmentation in the chimaera. From this, it was deduced that cell proportions in the coat were closely correlated with cell proportions in the growth controlling tissue or tissues. And that, as we were to discover later, was the main cause of the trouble. In retrospect, it is easy to see that we should have been alerted to the trouble at that point.

The next paper, Paper 16, is more cautious and was drafted in a more realistic manner. It reflects some awakening to the nature of the problem. By now, we had added eight organs to the coat for measuring the extent of the chimaerism. This was done by using two variants of an enzyme which could be spearated and quantified, the two component strains being marked with different variants. The linear relationship of body weight with cell proportions still held. But the proportions in each of the nine organs was correlated with size, while none of the nine predominated. Further, all nine taken together accounted for all the chimaeric variance in body weight. The paper puts it slightly differently, but the main conclusion is now inescapable : either these nine organs, jointly, controlled growth between them, or else the nine taken together accurately reflect the cellular genotype of the notional growth-controlling tissue. The paper comes close to admitting that there may be no such tissue, though it cunningly states this in the conditional mood. The alternative is stated quite explicitly. In the absence of a growth controlling organ, body size depends on cellular genotype throughout the whole body. Quite clearly, we had by now made a significant retreat from our earlier assumptions.

The definitive account of the programme comes in Paper 17. The account is complicated, and I cannot help feeling that it is unnecessarily so, but even yet I cannot see a ready way of simplifying it. By now, we had more chimaeras : 63 overt ones, which do not include some represented in earlier accounts. We had also added two more organs. A great deal of detail was filled in, and some earlier tentative conclusions were substantiated. Some

other preliminary conclusions were discreetly revised, but these were very minor issues. Truth be told, no major new conclusion emerged, except that possibly the weights of some organs were marginally affected by their own cell proportions. Despite this, and logically in contradiction to it, we were driven to the conclusion that growth regulation was systemic. In other words, we proceeded to dismiss possible local effects on organ size. Had the claim of local effects been stronger, editors and referees would no doubt have siezed on the flaw in logic, which only now has become apparent to me. But I do not wish to labour the point, because I feel that our main conclusion was correct. Given the systematic effect, it was not possible to decide whether this effect derived from some organ not studied or whether it derived, in some undefined way, from cell proportions in the body as a whole.

On a purely personal basis, and I emphasis the personal aspect to exempt my co-authors should they wish to differ, I now believe with the advantage of hindsight that the key issue was lost in the body of the paper. This was that chimaeras were incapable of answering the questions we set ourselves, and it would have been a service if the paper had highlighted the reason why. Among the organs studied, we included pituitaries and adrenals which were so small, physically, as to be at the limit of our technique for estimating the amount of chimaerism in them. Yet the correlations of these small organs with cell proportions in other tissues, and with body weight, were every bit as good as those of larger organs, like coat or liver. In other words, the degree of chimaerism is much too fine-grained to be of use for the purpose we had in mind. The cell proportions in the smallest organ is highly correlated with the proportions in the largest, and with the proportions in the body as a whole. This means that even the smallest organs do not derive from a small number of progenitor cells. Yet, given that we also found differences in cell proportions between organs within individuals, different organs presumably do not derive from the same mix of progenitor cells. All this must be highly relevant to questions of development and of early embryology, but by the nature of the system, chimaeras cannot contribute much to a study of the control of growth. That is where our initial assumptions went wrong, and that is the reason why a laborious and extensive study led to a slightly vague conclusion. This

final outcome was, I must admit, rather disappointing and yet, in an odd way, quite illuminating.

The next study to be reported was also collaborative but technically less ambitious than the chimaera one. It addressed what we regarded, at least initially, as a fundamental issue in the control of growth. This may be summarised as follows. If an animal is to become larger, it must do this by increasing the number of cells in its body, or by increasing the size of each cell, or by some multiplicative combination of the two effects. Conversely, smaller animals must have fewer and/or smaller cells. Now, an animal can become larger in one of two ways. It can either grow normally as it ages, or it can be selected, as is well known, for weight at a given age. We were particularly interested in the effects of selection on cell number and cell size, but we also examined the effects of ageing on the cell measurements and, as it turned out, it was just as well that we did.

The results are reported in Paper 18. Six replicates each were available of selected large strains, small strains and unselected controls. Counts of nuclei in each of four organs were used to estimate cell number, and we also derived estimates of individual cell mass, though the two estimates were not independent. The replication was an important feature of the study. If an effect was found in all six large strains, say, then a direct consequence of the selection for growth could be claimed with some confidence. If the effects were irregular and found only in some strains, then random drift, perhaps in an unrelated character, could be the cause. In the event the results, though not absolutely regular, were sufficiently so to allow some very clear conclusions to be drawn. The main one was that at six weeks of age, which had been the age when the mice were selected, large mice had more cells and larger cells than the controls, the small mice fewer cells and smaller. Thus, both cell number and cell size had been changed as a result of the selection, though by unequal amounts in different organs. In the case of the lung and the spleen, some 70% of the differences in organ weight were due to differences in cell number. The remaining 30%, by our definition of cell mass, had to be due to cell size. But in the case of the liver and kidney, the relative contributions of cell number and cell size were about equal. This might suggest that growth regulation operates through both cell number and cell size, but not in the same way throughout the body. Had it been left at that, a conclusion in those terms might have

raised more questions that it answered. Fortunately, we did not have to leave it at that, and a subsidiary study on changes in cellular properties as mice grew older proved crucial to the interpretation.

The subsidiary study had been based on fewer mice than the main experiment, and the results in consequence were somewhat more erratic. Nevertheless, their interpretation was still clear. We took the liberty of excluding some troublesome data on livers beyond six weeks of age, with only a hint that polyploidy (for which the organ is noted) might invalidate the data from the livers of older mice. With that exception, the pattern was quite distinct. As the mice grew older, it was shown that cell numbers and cell size both contributed to increases in organ weights. Furthermore, the proportionate contributions were very much the same as those found for the selected strains. Put crudely, this meant that given a named organ of a given weight, we could retrospectively predict both its cell number and its mean cell mass, irrespective of either the age of the mouse or the strain from which it came. Actually, its sex might have made a small difference, but that effect was too trivial to invalidate the statement just made.

The final conclusion of the paper, taking the main and subsidiary experiments together, was that selection did not have separate effects on cell number and cell size, but that both were predictably affected by the speeding up or slowing down the normal processes of cellular growth. A mouse selected for high six week weight is no different, as far as its cells are concerned, from what a small mouse would be at a much later age when, or if, it reached that weight. We began by asking how growth was affected by cell number and cell size. The final answer was that growth was not affected by the cellular properties but rather, that the cellular properties were themselves affected by growth. The paper shuns putting the statement quite as baldly as that, but it ends with the suggestion that this kind of conclusion has a generality far beyond the cellular context in which it was framed. Growth seems to act as a determinant, and the rest follow.

The next study on the genetic control of growth returns to the topic of maternal effects, which were discussed earlier in Paper 11. There, it was shown that maternal effects is our material were stronger postnatally than prenatally, but the pattern was asymmetrical. Small mothers proved inadequate for large offspring, whereas small offspring failed to profit from

large mothers. Paper 19 examines these postnatal effect in more detail. Despite endeavours to avoid the fault at the time, this paper to my mind still suffers from an excess of numerical data. It could not have been an easy account to read for any reader lacking total commitment, and this probably explains why it has not been widely quoted. This is a pity, because the paper has some implications. Briefly, it compares the relative importance of maternal effects and genetic background on adiposity (fatness might have been a better term). An elegant crossfostering design was adopted, and I may be excused the adjective because the details of the design belonged to my coauthor. Equal numbers of large and small offspring were reared by both large and small dams in standardised litters of 4 and 8. Statistically, the design was efficient and powerful, as reflected by the clarity of the results.

Both a large genotype and a large foster mother led to fatter mice at six weeks of age, the direct genetic effects being about three times more important than the maternal effects. An interesting interaction between genotype and line of foster dam was generated by the inability of small dams to provide enough milk for large offspring, which made those large offspring less fat than they would be otherwise. In the same way, litters of eight were less fat than litters of four. But all this is at a constant age. At the same weight, small mice were always slightly fatter than large mice, and neither the line of the foster dam nor the size of the litter in which the mouse was reared seemed to make any difference to this general conclusion. In other words, when compared at the same weight, small mice are more mature than large mice.

The paper adds to our understanding of fat development in large and small mice, and of the factors affecting it. This brief review has been limited to fat, but several other metric measurements were reported, all with similar findings. My main reaction to reviewing this paper and looking again at the results is one of regret that I can not rewrite it. I think I could now do a better job. But the work still contributes to our understanding of the relationship between body size and fatness.

To the extent that I am right to say that the presentation of this paper was too numerical, the next one perhaps goes too far the other way. It condenses

a vast amount of numerical information into a few simple figures, and as it describes three different experiments, it might perhaps have appeared as three separate papers. My excuse at the time was that some of the internal coherence of the story might have been lost by separate publication. Be that as it may, Paper 20 describes a series of experiments on the effects of selection for body weight on food intake and the efficiency of conversion. As it turned out, the studies possibly told us more about the genetic control of food intake and efficiency than about the consquences of the original selection for weight. Whether that is correct or not, the results illustrate the power of selected material to explore wider biological phenomena.

The first study in the paper examined the food intake and efficiency of large and small strains, and their unselected controls, from weaning time to ripe old age. As far as I am aware, the data are unique in terms of the age span covered, and they allow an unambiguous measurement of the maintenance requirements of mature mice. Maintenance requirements of mature mice were related to metabolic body weights, and there is no surprise about that. It was not a novel finding either that large mice both ate more and converted it more efficiently than small mice, when they were compared at the same age. But it was novel to find that this was also true, for both intake and efficiency, when the strains were compared at the same weight. In the paper, I deal with food intake in terms of differences generated by selection, where the large mice have a curious peak in their intake at around six weeks of age. They then reduce their intake to a steady state at a lower level. I seem to have dismissed their greater efficiency as a function of this increased intake. Looking again at the data, I seem to have missed something quite important. If we compare the strains at a body weight of, say, 20 grammes, the small strains need virtually the whole of their intake for maintenance. The large mice at this weight will eat about 7 grammes more food over the next week and gain about 10 grammes in body weight. It is thus clear that differences in efficiency can not be related in a simple manner to differences in intake or in metabolic weight. Selection for increased body weight therefore not only increased appetite but brought about real changes in metabolism, possibly related to a lower rate of protein turnover and thus in maintenance requirements.

The next study was designed to explore differences in the mechanism of

appetite control, but did not work according to plan. It did, however, establish the following. When some of the energy was fed as a glucose solution, the intake of solid food was reduced in exact energetic proportion. Further, growth was not affected when glucose was fed. Since solid food contains protein and other nutritional requirements, these substances could not have been limiting in the circumstances of these experiments. The conclusion is clear : appetite control, or at least short-term appetite control, is regulated by energy requirements. It also seemed that large mice had a preference for glucose whereas small mice did not. While I am not too confident about that statement, it suggests that, if it is true, the greater energy requirements of the large strains made them opt for the ready source.

The third study reported in the paper identifies also a long-term appetite control, though it does not follow that the mechanism is the same. My guess is that it might be different. Unfortunately, and I now regret it, the only comparisons were between large strains and their controls; the study would have been more powerful if the small strains had been included. Briefly, the study shows that following a period of food restriction, animals on resumption of full feeding eat according to age but convert it according to weight. Thus, a mouse formerly restricted would, on resumption of full feeding, eat as much as an unrestricted mouse of the same age, which was perhaps twice its size. But it would convert that intake with the efficiency of a mouse of its own weight that would, normally, be much younger. Compensatory growth is thus the product of this high intake and high efficiency. I believe that this may be the best way to describe compensatory growth, though obviously it provides nothing by way of explanation.

Dr John McCarthy (personal communication) was able to confirm my result, and happily removed a possible criticism of it, This was that restricted mice were simply hungry and that they merely restored gut-fill to normal levels. This certainly occurred by a conspicuous "overshoot" of intake over two days or so. Thereafter, his results match mine exactly, and the conclusions stand.

Long-term appetite control is therefore governed by some kind of clock that is not related to the animal's growth. The animal's efficiency, on the other

hand, can be accurately predicted by its weight, which is probably a reflection of its developmental stage. But weight is not the sole determinant. Large mice were more efficient than their controls at all weights. So when an animal adopts an efficiency appropriate to its weight, it is the weight for that strain that counts. Thus, even a fundamental connection between weight and efficiency can be altered by selection, which therefore shows that efficiency has a genetic basis which is at least partly independent of that for body weight.

I have reviewed this paper at quite some length, because I regard it as my most important piece of experimental work. At least it helped me to think about the whole package of growth, feed intake, efficiency and body composition, though I addressed the last of these only by implication. What I should like to think is that the paper might help in the design of meaningful experiments in this area.

That was really as far as I went in examining the biological basis of the changes in growth brought about by selection. The remaining two papers in this section perhaps should not have been included, but both arose out of selected material and, initially at least, it was possible that they were some side-effect of the growth control. In the event, it seems that they were not, but negative results should also be reported. In any case, they illustrate how genetic interests can flow naturally from one area into another.

Paper 21 describes a case of polydactyly that arose spontaneously in a line selected for large size. It proved to be an excellent example of a condition that was familially transmitted but which failed to reveal any acceptable segregation ratio. Its genetic interpretation was eventually resolved in terms of Falconer's threshold model, with an underlying continuous distribution of a notional liability. After several false starts and wasted effort, the trick in the analysis was to find a classification of the multifaceted data that rendered this interpretation reasonable. The main features were that two thresholds were established, one for hind foot polydactyly and a second, more severe condition, that affected the fore feet as well. The syndrome proved to be complex, with various internal correlations and random effects governing the phenotypic manifestations. It is reasonable to suppose that a major gene was involved

but with so many modifiers that segregation could not be observed, despite my best efforts. There were interactions between local and systemic effects on digit number, with implications for the genetic control of limb development. But to the best of my knowledge, these implications were not taken up by anyone, and the paper seems to have had little impact.

Another condition, described in Paper 22, also arose in a large line, and this proved to be a straightforward dominant gene, lethal when homozygous. It was called 'Small eyes', which describes the condition perfectly. It proved not to be allelic with another mutant called microphthalmia. The new mutant was subjected to extensive examination by Dr R M Clayton and her collaborators, who concluded that the basic lesion was a membrane deficiency that affected transport across it. It was rewarding to see the mutant proving to be such good research material in the hands of colleagues.

Group IV. Non-experimental studies

I have never done any theoretical work worthy of the name, yet it could well be true that some of my best work has been non-experimental. I suppose I have contributed my share of review articles, usually by invitation or perhaps to secure my passage to some international gathering. I should like to think, however, that in the main I was not content on these occasions merely to review a topic. Instead, I attempted to use the published records in a more constructive manner, to establish ideas and clarify the thinking on the subject, or at least my thinking. On occasions, I sought to develop new presentations of material familiar to quantitative geneticists, to make it more accessible to a wider audience, perhaps from other disciplines. In particular, I was frequently concerned to translate the results of experimental work with the laboratory mouse into an animal breeding context. As I said earlier this was the motiviation for much of my own experimental work with the mouse and I felt it was important to me, from time to time, to try to put it in perspective, and attempt to assess its value.

This section reviews a sample of such articles. I have omitted some of my more pedestrian attempts, and also some others that are similar to those included. I shall merely attempt to assess the significance of each contribution without going into too much detail on content.

The first two papers in this section can conveniently be taken together. Both concern the application of work on the laboratory mouse to animal breeding research. Paper 23 refers to the genetics of body weight, while Paper 24 deals with all other traits but mostly litter size and fertility. Both were published in successive numbers of the same journal, but they are essentially two parts of the one review. The date was 1965, and this was possibly the last chance for any one reviewer to cover the field comprehensively; even then the review was too long to be published without being split. It was just before some creative workers, like Eisen, Bradford, McCarthy and Sutherland - some of whom have since published extensively - had made their marks. At the time, it was just about possible to hold all the relevant literature in the head long enough to write the review. Later, as I was to discover, this became a hopeless task, though ageing on my part may have had something to do with it.

Two main conclusions of the review now strikes me as interesting. One was that elaborate and costly schemes to exploit nonadditive genetic variance had not really worked for any trait, not even for litter size. Straightforward selection had always worked better. This was before Falconer's development of inbreeding and crossing to rid selected lines of deleterious recessives, reviewed earlier in conjunction with Paper 10. However, that was a special case. The other main conclusion was that interactions, by and large, were singularly unimportant. Generally, I believe that these early conclusions still hold. Some areas, where I claimed we were ignorant but where I must have known what was in progress, have since been well documented, like the repeatability of selection responses and their limits. I also called for more experimental work on the effect of population size on responses and on the balance between intensity of selection and inbreeding. If I meant it seriously, it was naive to expect this work to be done on mammals. We now probably know enough about these topics from theory and from experimental work on Drosophila, but they do indicate some of the concerns of the time.

The main value of the review was probably its comprehensiveness, and the lead it gave into the existing literature. But despite my comparative inexperience, or perhaps because of it, I seem to have taken it upon myself to write a highly evaluative report of the state of the art at the time.

The review certainly had an effect, and for some years, I was disappointed if I came across a paper on mice, in the area, that did not quote my review! To my astonishment, I find that it is still being quoted in 1984. But I am no longer gratified to see it quoted. This is because the review is not being used any more for its original purpose, but to extract statements, usually out of context, which the author finds convenient to support his case. There comes a point when review articles should be quoted solely in the context of their time, to illustrate historical viewpoints.

The next paper, Paper 25, is in my own opinion my best piece of writing, ever. The material is not in the least bit original, and relies totally on Falconer's Introduction to Quantitative Genetics. The first edition of that book had been in existence for about two years before my article was written, though in the event it took another five years before the volume for which the article was destined became published. I may have contributed marginally to the delay in publication by entering into a fierce argument with McGraw-Hill, the publishers; I insisted that the plural form of formula was formulae, and not formulas! I rather doubt by now whether this kind of defence of linguistic practice is worthwhile for technical articles, particularly as I was prepared to accept American spelling for other words.

My brief for this article was a simple one : I had to present the concepts and methodology of quantitative genetics to a lay audience with no training, in this case behavioural scientists. To be fair, I was allowed to assume a reasonable grasp of statistics, and without that, the task would have been impossible. Falconer's book had opened up new vistas for many of us, and a depth of understanding we had not experienced earlier. Even so, the book was still too sophisticated for the uninitiated, and in any case too long. All I had to build on was a potential interest, and my task was to encourage my readership to seek further. I had to give them the confidence, for a start, and then to show the relevance. Some of the latter was done in other chapters in the same volume, not included in this collection. Paper 25 was my effort to teach them, and shedding all modesty aside, I seem to have succeeded beyond my expectation. The chapter became the core for several university courses, and quite literally launched some people on their careers, if I am to believe their testimony. My chapter made Falconer's book accessible to them, and from there, several went much

further.

Looking at the chapter again, I am a bit mystified as to why it was so successful, particularly as it ignores population genetics and goes straight into the quantitative. I can only imagine that because, at the time, I had only newly mastered the material myself, I was probably more alert to its potential obscurities than I might have been subsequently.

Paper 26 also influenced the same audience, but in this case, I am not nearly so proud of the fact. The paper purports to trace the evolutionary significance of some behavioural traits, by examining their current genetical structure. The general thesis is that traits close to natural fitness display little additive variance, but will still display nonadditive genetic variance. This had been known to all geneticists since Fisher's Fundamental Theorem of Natural Selection, and I seem to have made an inordinate meal of saying so. And having said so, I then went on to develop the alternative viewpoint, by adapting a point originally developed by Alan Robertson. That was that behavioural traits, while possibly connected with fitness, can still display considerable additive variance because natural selection might have been for an intermediate optimum.

The paper, frankly, is weak on theory and uncritical of experimental evidence. From both points of view, the least said about the better. I have included it only because it has had one saving grace. It has made people think who otherwise might not have thought. It made some people aware of the connection between evolution and current genetic features, and how animal behaviour may possibly fit into the pattern. With all its weaknesses, the paper has given behaviourists a basis for orderly thought where previously they were muddled. They now at least realise the possibility that things are as they are because of what may have happened in the past. They may have come to realise this anyway, but judging from the number of references to it, my paper played its part in bringing about this realisation. I am more than a little embarrassed to find it still being quoted - at least twice so far in 1984 - but if people still find the ideas useful, the paper at least may have said something they wanted to hear.

Paper 27 discusses the side effects of selection for growth, with some emphasis on the application to animal breeding. Correlated responses had often been discussed before, but largely within the customary biometrical framework with emphasis on genetic correlations and the like. In this paper, I attempted to interpret the findings in a different way, by discussing correlated responses in terms of the animal's biology. The conclusion, based on a limited range of experiments and largely my own, was that some of the deleterious effects were a direct consequence of large animals being fat. Fertility, for instance, could be restored either by restricting an animal's food intake or by mating it at an earlier age, before fat had accumulated. Such animals were still genetically large, and to some degree physically large, and yet would breed normally. Thus, it is not the direct effects of genes for large size that themselves cause the fertility problem, but rather the internal environment they provide for the animal's other physiological functions. The usual biometrical approach does not make this distinction and the effects on fertility fall neatly into the category of pleiotropic effects. But to accommodate the distinction I drew, I invented the phrase "pleiotrophy once removed"; not surprisingly, and perhaps thankfully, the phrase did not catch on.

The paper would have been better organised had I, as I have done here, dealt with fertility first, where the point can be made clearly. The same sort of conclusion in effect applies to body composition. Large animals are fatter, especially as they grow older. This, I argued, was another property of the animal's internal environment. Selection for large size often does not make the animal fatter at the age when it is selected. But because an increased food intake is necessary to achieve the weight given to that age, the food intake is excessive as growth begins to asymptote. The consequence is that fat is laid down. But I argue further that this is not an inevitable consequence, and a few experiments bear this out. There does seem to be genetic variation in the partitioning of energetic input, especially into fat and heat. This explains why fat animals are often more efficient, and the simple explanation is that it is more efficient to lay down fat than it is to lay down nothing. I bolster my argument in the paper by applying some bioenergetic considerations, particularly Webster's work, to the interpretation of selection responses. By and large, the internal coherence of the arguments still convinces me.

At the end of the day, I do not think it matters much whether my arguments are right or wrong. The paper has had some effect, and has been widely quoted. Its value, such as it is, was to encourage a way of looking at correlated responses in a less stereotyped fashion, and to seek a more biological interpretation of genetic correlations. This in no way invalidates the biometrical description of a system, but perhaps it prompts a more illuminating way of assessing the consequences. The paper is not particularly well written, but in terms of content, I am still reasonably satisfied with it.

Paper 28 considers further the genetical control of growth and fertility, and seeks to rationalise a mass of laboratory evidence in evolutionary terms. It was not a successful attempt, and my efforts to be original were rather too obviously contrived. I do however extend the usual discussion of laboratory experiments to studies on wild mice. There is not a lot of evidence on wild populations, but I might have made more of one very clear conclusion. Reading the paper again, I am struck that when wild mice are brought into the laboratory, their genetical properties are remarkably similar to those of laboratory mice. This presumably implies that the effects of natural selection in the wild are not grossly dissimilar to what happens in the laboratory, and this I should not have expected. Perhaps it is partly a matter of time scale, and that the laboratory mouse, in evolutionary terms, is not that far removed from its wild ancestors. Be that as it may, my main conclusion was that both body weight and litter size had intermediate optima, in terms of fitness. There may be a range of intermediate values, for both traits, over which fitness is not much affected. I do make the point in the paper that this statement, as it stands, is a trite observation. It is probably true of all traits at all times; the real issue is whether there is a wider zone of neutrality for some traits than for others. The answer given in the paper was not entirely original. It was that the best guide to the width of the zone of neutrality is given by the heritability of the trait. If that is right, body weight is more tolerant of deviations around the mean than litter size, in terms of the animal's fitness.

The two most interesting conclusions in the paper, or perhaps I should say

speculations, are lost in the body of the text. The first is that the high degree of embryonic mortality, usually found both in the laboratory and in the wild, seems to have little if any genetic basis and consequently must be environmentally induced. But since embryonic loss has a non-zero energetic cost, if nothing else, why has it become an established feature of mouse populations? My suggested answer in the paper is that it is a form of evolutionary insurance, so that the mouse can respond rapidly to temporary improvements in the environment, such as a good season or perhaps an unexpectedly provident niche. To exploit such circumstances, some embryonic mortality under normal conditions might be a reasonable premium to pay.

My second speculation could be of some biological significance. Reviewing the evidence on growth control, I note the well-known phenomenon of compensatory growth, where animals can recover rapidly after long periods of privation. This must be very important in the wild. I illustrate this with the little-known case of desert rodents that are chronically hyperphagic, given the chance. After some rain, these rodents rapidly accumulate large stores of fat, so much so that they can survive 30 days of starvation, whereas a normal mouse would be dead in two days. Growth control must therefore ignore ephemeral perturbations, which may be at least one reason why the growth control system is so intractable to experimental attack. It makes me wonder now how much effort on growth control I might have saved myself had I realised the main feature of the system in time.

The next paper I shall always regard fondly, for the simple reason that I got to deliver it on a memorable visit to New Zealand. It was also a swansong, because it happened at the time when I had to give up work on laboratory mice. Paper 29 returns to the application of laboratory experiments on the mouse to animal breeding research, and it was to be the last time I had anything to say on the subject. That was maybe just as well. It was not the first time (or the second or the third) for me to be asked to address this topic, and the paper shows clearly how I was struggling for something new to say. My solution was to become a little more philosophical than had previously been my wont, or at least my practice. I did manage to draw some general conclusion on the effective population size needed for a practical breeding programme and I discussed the reliability and predictability of selection responses, with experimental illustrations. The main

conclusion was that the real value of laboratory experiments was to explore concepts and ideas, leading to better understanding of the ingredients of a practical breeding programme. This can not normally be done within the constraints of multi-trait selection on large animals and we need further analyses of the biological basis of a range of complex phenotypes. I am not sure that this was really worth saying. It is the kind of sentiment to which most people would subscribe. The paper is included only because it represented my assessment, after more than twenty years in the field, of how laboratory animals should fit into the general area of animal breeding research.

The last paper in the collection might perhaps have been omitted, because it changes key. But I have included it because I believe it bridges the previous twenty nine papers with what are likely to be the more urgent needs of animal breeding in the future. Animal improvement in the past has leant heavily on the concepts and methods of quantitative genetics, with its assumption of many genes with small effects. And there is no reason to doubt that this methodology will still be the main source of improvement for some time to come, at least for most species. Paper 30, however, examines the exploitation of genes with large effects in animal breeding programmes. Several examples quoted in the paper suggest that such genes may already be of increasing importance. Further, the rapid advances in molecular biology have now made gene transfer a practical proposition and it may soon become routine. If nothing else, the molecular techniques will very soon allow us to dissect the genome into discrete entities; their manipulation must only be a matter of time. The paper discusses how the new information can be exploited in animal breeding, and examines the consequences of genes with large effects on the standard parameters like heritability and the genetic correlation. It also warns against a kind of statistical naïveté in handling this new information. In short, it points out the need for a full economic assessment of the effects of the gene on every production trait for each of the three genotypes at a locus. As experience with the halothane locus in pigs has shown, this is no mean task, but the dangers of proceeding in the absence of all the information could mean that a selection programme might be seriously misdirected. The paper, on reading it again, seems to strike a realistic balance between the opportunities that the new techniques offer and the limitation of their practical application.

That completes the review of the papers. They have spanned a time of change. Some of the early papers already seem dated, while the last of them looks at future possibilities. Animal breeding faces new challenges as the production of most animal commodities is now in surplus, at least in terms of the markets of the developed world. As a science, animal breeding has drawn on its parent discipline of quantitative genetics and has played its part in removing the food shortages that have existed within living memory. There is noreason to suppoose that the story will end here. Animal breeders of the future will continue to develop their products in response to consumer needs, and to the extent that new developments in genetics can help, they will no doubt be adopted and exploited.

STATEMENT OF AUTHORSHIP

I am the sole author of 16 of the 30 papers in this collection, where the work was my own except for technical help and the benefit of the usual interactions, on a reciprocal basis, with friends and colleagues. For two of these papers (1 and 14) the data had been collected for other purposes by Professor D S Falconer, but the treatment of these data for the publications included here was my own. For the papers reported under joint authorship, the work was in all cases closely collaborative and even at the time, individual contributions could not be reliably quantified. I have nevertheless attempted to indicate my contributions, under three headings :

initiation - developing the ideas and experimental design

execution - including the organisation of genetic material, the supervision and motivation of staff

completion - data reduction, analyses and publication

I can only hope that my recollection of my contribution to these papers is reasonably accurate.

	% contribution to		
Paper	Initiation	Execution	Completion
2	40	40	30
3	20	20	15
4	95	50	20
10	95	25	75
11	95	25	75
12	80	25	40
13	100	100	45
15	20	20	60
16	20	20	10
17	20	20	5
18	20	25	20
19	10	5	40
21	100	100	60
30	10	N/A	80

The work included has not been submitted for other degrees, with the following exceptions. Paper 1 was based on part of my Ph.D thesis (University of Edinburgh, 1956), while parts of the same thesis contributed data to Papers 2, 3 and 4. Papers 10 and 11 were based on parts of a Ph.D

thesis by W K Al-Murrani (University of Edinburgh, 1973), while Paper 12 was based on part of a Ph.D thesis by E Yuksel (University of Edinburgh, 1978). The data for Paper 13 were used in part for an M.Sc. thesis by D E Steane (University of Edinburgh, 1979).

ACKNOWLEDGEMENTS

Almost all the work reported in this collection was conducted at the Institute of Animal Genetics, Edinburgh. I look back with gratitude on the privilege of working there, and it is a **pleasure** to acknowledge the benefit and stimulation I derived from colleagues and the many visitors to the Institute. Among my colleagues, I am particularly indebted to Professor D S Falconer, who was my closest collaborator, and who was ever generous in his interest and support. I was also able to profit from innumerable discussions with Professors Alan Robertson and W G Hill, which I value beyond words. For many years, Dr I K Gauld was a congenial and supportive colleague, and it was only through his skill and perseverance that some of our joint work was brought to a successful conclusion. Finally, I deeply appreciate the loyal support provided, over different periods, by Mr J H Isaacson, Miss H I Macrae and Miss A Walker, and their staffs in the mouse house; without their meticulous technical work, much of the work reported here would not have been possible.

PAPER 1

The effects on litter size of crossing lines of mice inbred without selection

Genet. Res. 1, 239 - 252. 1960.

by

R.C. ROBERTS

Genet. Res., Camb. (1960), **1**, *pp.* 239-252
With 3 text-figures
Printed in Great Britain

The effects on litter size of crossing lines of mice inbred without selection

By R. C. ROBERTS
Agricultural Research Council Unit of Animal Genetics, Institute of Animal Genetics, Edinburgh 9
Received 2 October 1959

1. INTRODUCTION

In the exploitation of heterosis in plant and animal improvement, inbreeding is frequently employed to produce genetic differentiation in the parent material, through random changes in gene frequencies. But if the resultant crosses are to represent genetic progress, random differentiation alone becomes insufficient, for genotypes of special merit cannot be provided without selection at some stage. The selection, which may be either natural or artificial, can apply or be applied at any of these stages:

1. Selection within lines on inbred performance.
2. Selection between lines on inbred performance.
3. Selection between lines on crossing performance (general combining ability).
4. Selection between crosses on cross performance (special combining ability).

Natural selection will act primarily through the first two ways, if the character is an aspect of natural fitness. Artificial selection can also be applied during inbreeding, though theoretically it is best reserved until the crossing programme.

Little, however, is known about the efficacy of selection in the context of inbreeding, and the experiment to be described here was designed to provide information relevant to this general problem. The character studied was litter size in the mouse. The results are therefore relevant to some problems in animal improvement, especially to such characters as the fertility of pigs.

The general plan was to inbreed a number of lines without any artificial selection, and with minimal natural selection. To this end, it was imperative to preserve all possible lines. This in turn precluded raising the inbreeding coefficient above 50% or so, for by previous experience the loss of lines then becomes inevitable. Thus the experiment was of necessity restricted to only partly inbred material, but the obvious theoretical disadvantage of this was somewhat mitigated by greater practical application. For the difficulty and cost of maintaining inbred lines becomes prohibitive in farm animals, even in pigs (see, for instance, Donald, 1955), so that the use of partly inbred material must be explored.

The lines were crossed to obtain the following information:

(*a*) To compare the performance of the crossbred population with that of the original outbred population from which the inbreds were derived. This comparison would indicate what improvement, if any, would accrue from natural selection which operated almost entirely within lines (1 above).

(*b*) To estimate the variances of general and special combining abilities. These estimates would assess the effect of artificial selection applied in the manner of 3 and 4 above.

The application of selection in the manner of 1 and 2 above was the subject of another experiment, on the same stock of mice, described by Bowman and Falconer (1960).

2. THE CHARACTER—LITTER SIZE

Litter size would appear to be a self-explanatory term—the number of young born in a litter. This definition is unfortunately complicated by the disposition of mice to eat many of their still-born young—and possibly some others as well. The number of young found is thus influenced by the interval between birth and the examination of the litter. In the experiment reported here, cages were examined once daily, the number of live young being recorded as the litter size.

All the work was done on first litters only. The collection of sufficient information on second litters to be of material assistance would inordinately prolong the generation interval, sufficiently so as to nullify the advantage of more accurate measurement. The character chosen for study was therefore 'the number of live young found in the first litter'. While this may not reflect accurately the common concept of 'litter size', the term as defined has complete operational validity.

Litter size as a character is one of considerable complexity. It has three major factors, each of which determines the upper limit of the succeeding one:

1. The number of ova shed.
2. The number of ova fertilized.
3. The number of zygotes carried to term.

The first of these is of course wholly a character of the dam. The second may be influenced by either the sire or the dam. Though Falconer (1955) showed the effect of the sire on litter size in outbreds to be negligible, this may not be so in an inbred population. While it may be tempting to regard the third component as a function of the viability of the young, we cannot exclude the potential influence of the dam, quite apart from her contribution to the genotype of the litter. It can be seen, therefore, that when litter size as a character is submitted to any genetical analysis, its constituent factors are intricately confounded. This problem will be discussed at greater length when the actual results are examined.

The complexity of the character, however, does not end with its multiple determination. For the number of young born is subject to a strong maternal effect dependent upon the weight of the mother. A large mother tends to produce a large litter, in which individual weights are consequently depressed. This handicap is still reflected in weight at mating time. Hence the daughter of a large mother tends to be light, and produces a small litter when she in turn bears offspring. The net effect is thus a negative regression of litter size on the size of the litter in which the dam was born, unless there also exists the positive genetic pathway expected of a heritable character. These complicated interactions were studied by Falconer (1955), who calculated the path coefficients relating litter size to the body weight of

the dam and the size of the litter in which the dam was born. The path diagram is shown in Fig. 1. The mother's body weight is inversely correlated with the size of the litter in which she was born, and directly with the size of her own litter. The product of these two coefficients is $-0{\cdot}07$, which would give the regression coefficient of litter size on maternal litter size if no other pathway were operative. There is, however, a direct genetic pathway, which is measured as the partial regression of litter size on maternal litter size holding the mother's weight constant. This coefficient is $+0{\cdot}07$, as shown. From this, we see that litter size is affected by maternal

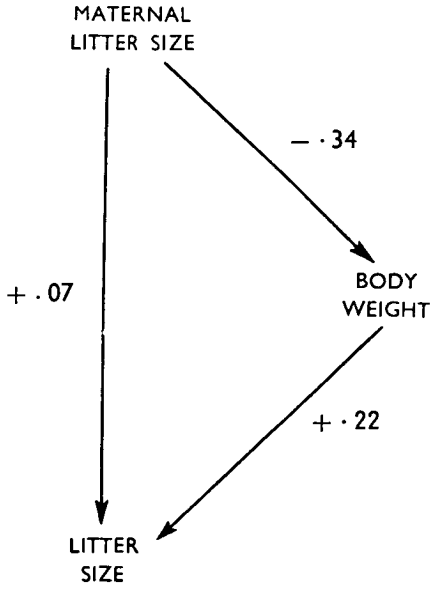


Fig. 1. Path diagram and standardized partial regression coefficients. After Falconer (1955).

litter size through two independent pathways of equal magnitude but opposite sign, explaining why the direct regression, when measured, comes out to be zero (Falconer, 1955).

The relevance of these maternal effects to the subject of this paper is apparent from the following considerations. Firstly, as litter size declines on inbreeding, a maternal effect will be initiated counteracting to some extent the direct effect of inbreeding. But body size itself, and other characters such as milk production, are also liable to be depressed by inbreeding. The possible interactions become so involved that the nett effect becomes obscure. At the present state of our knowledge, we can only approach the problem empirically, and this discussion of the complexity of the character has been presented to show that any attempts at a more sophisticated interpretation of the subject would only be of questionable validity.

3. MATERIALS AND METHODS

It was a premise of the experiment that no lines should be lost during inbreeding. Previous experience had shown that in practice it would be impossible to carry the

inbreeding coefficient beyond about 0·50 without introducing the likelihood of losing lines through low fertility or even complete sterility. It was clear from the start that the crossing would have to be done from partly inbred material.

The broad outline of the experiment was therefore as follows. The inbreeding stage was confined to three generations of brother-sister mating. The lines were then crossed at random giving crossbred litters. As litter size is largely a maternal character, these crossbreds had to be mated to test their fertility, for this was what the experiment was required to determine.

(i) *Inbreeding programme*

It was decided to start with thirty inbred lines, which were derived from mice surplus to the requirements of a selection experiment for litter size described by Falconer (1955). Ten inbred lines were derived from each of the high, low and control stocks of the selection experiment, which had then proceeded for ten generations. There was therefore some differentiation among the original material which had to be allowed for in the crossing programme. Ten litters were chosen from each stock; each litter came from one family and subsequently became the foundation of one inbred line. The largest and smallest litters in the 'high' and 'low' stocks respectively were of course required for the selection experiment. In choosing litters for the present work, this bias was counteracted by rejecting also the other extreme. With this exception, and the avoidance of sib litters, the foundation litters were taken at random.

The inbred lines were propagated in the following manner. All the available females of a litter were divided between two of their sib males, as a precaution against male sterility or accidental loss. Each line thus normally gave birth to more than one litter, one of which was taken at random. The random choice was occasionally disturbed by a litter not containing the required two males and two females, which was usually excluded in the interest of safeguarding the line. But any selection thereby introduced against litters of extreme sex ratio and against some small litters was so slight (and probably ineffective) that it was considered to be of little consequence.

The mice were mated when the youngest reached 6 weeks of age, the oldest mice of that generation being approximately 8 weeks by that time.

In spite of all reasonable efforts to maintain them, four lines in fact failed to complete the inbreeding stage of the experiment, and of course are not represented in the crosses. Two lines were lost for reasons unconnected with fertility, but the loss of the other two must be ascribed at least in part to low fertility. Each gave birth to small litters, all of which died before weaning. There was therefore undoubtedly a little selection during inbreeding, but its magnitude must be considered insufficient to affect materially any conclusions that emerged from the work.

(ii) *Crossing programme*

Ideally, each line should be crossed to all the other lines to form an orthogonal set of diallel crosses, but this was prevented by the exigencies of space. Any system

whereby the crossing was done at random would meet the basic requirements of the experiment, and the principle of the scheme finally adopted is illustrated in Fig. 2. In order to use all available lines as both male and female parents, pair-matings were employed. The size of the litter of any one pair was an estimate of the value of that cross. A certain number of replicate crosses was therefore required to assess the error variance.

All crossing was done within each of the three major groups from which the inbreds were derived. The scheme depicted in Fig. 2 was therefore used for each

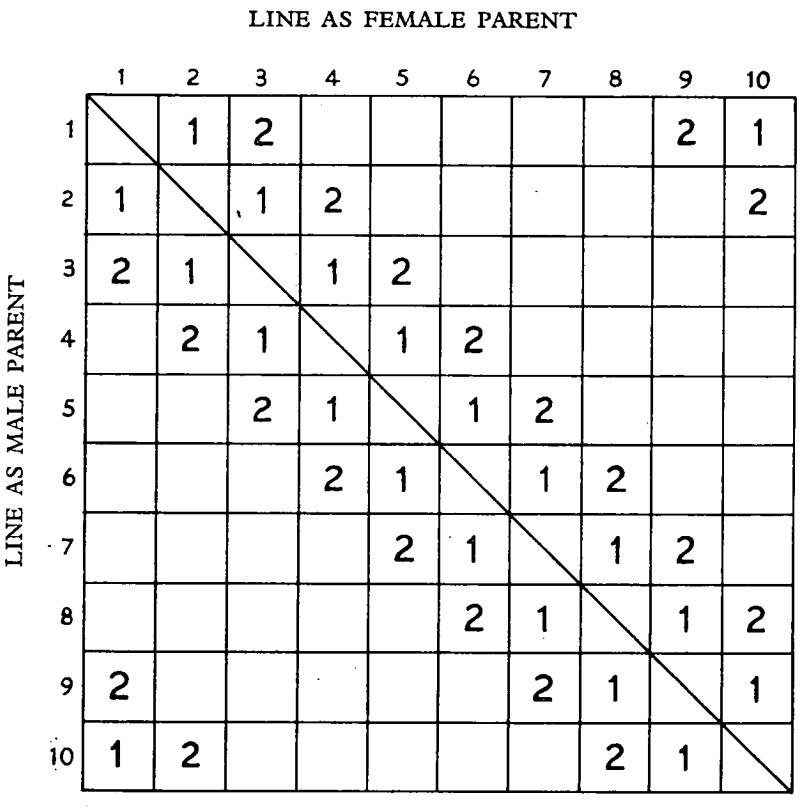


Fig. 2. The principle of the scheme of crossing the inbred lines. The number in each cell represents the number of matings between those lines.

group in turn. There were insufficient mice available to make all the matings required by this general scheme, but when a particular mating could not be made no other was substituted. This would introduce the least bias into the crossing programme.

The crossing programme required two stages, one to obtain the crossbred animals and another to test their fertility. The first cross measures the effect on litter size of crossbreeding in the litter, but still from an inbred mother. The second cross measures the further effect on litter size brought about by using a crossbred mother. In the second cross, litter size is regarded as a maternal character, as the direct effect of the male on litter size in fertile outbreds was known to be negligible (Falconer,

1955). We are interested in the effect on litter size of the genotype of the mother, as determined by her inbred parents.

In the second cross, matings between crossbred mice with a common parental line was avoided. Apart from that, the mice were mated schematically as before. The scheme of crossing employed was in principle a repeat of the first cross, except that a certain number of triplicate matings were substituted for the duplicate ones of the first cross. This was done as the error variance in the first cross was rather large.

As the inbreeding of parents and offspring are out of step throughout the experiment, Table 1 shows the inbreeding coefficients of parents and of offspring for every generation. The foundation animals are designated generation O, the inbred generations I, and the crosses X.

Table 1

Generation	Inbreeding coefficient		Litter size	Body weight of dam at 6 weeks (g.)
	Parents	Offspring		
O	0	0	8·12	21·9
I_1	0	0·25	6·73	21·2
I_2	0·25	0·375	5·82	20·8
I_3	0·375	0·50	5·69	20·1
X_1	0·50	0	6·20	21·5
X_2	0	0	8·47	21·3

4. RESULTS

The data that accrued from the experimental work will be presented in three sections in the following order:

(*a*) The effects of inbreeding and crossing on mean litter size.
(*b*) The differentiation between inbred lines in litter size.
(*c*) The analysis of variance of litter size in crosses between inbred lines.

(*a*) *Mean litter size*

To a limited extent, it is possible to observe separately the effect on litter size of inbreeding in the dam and inbreeding in the litter. In the first inbred generation, any reduction in litter size is clearly attributable to inbreeding in the young, as the parents are still outbred. Likewise, any increase in the first cross will be due to crossbreeding in the litter, and any further increase in the four-line crosses can be ascribed to crossbreeding in the parents. But, for the intermediate generations of the experiment, the inbreeding of parents and young will proceed simultaneously but at different stages.

The generation means for litter size during the inbreeding and crossing phases of the experiment are shown in Table 1, and are illustrated graphically in Fig. 3. The general picture is the expected one of decline on inbreeding, with subsequent recovery on crossing the inbred lines. In the first generation of inbreeding, mean litter size fell by 1·39 as a result of increasing the inbreeding coefficient of the young from

0 to 0·25. Over the next two generations, there was a further fall of 1·04 in mean litter size; as indicated earlier, it cannot be determined to what extent this is due to further inbreeding in the young, and to what extent it is caused by inbreeding in the parents.

In the first crossbred generation, when the inbreeding coefficient of the young was changed from 0·50 to 0, litter size improved by 0·51. This, of course, is a minimal

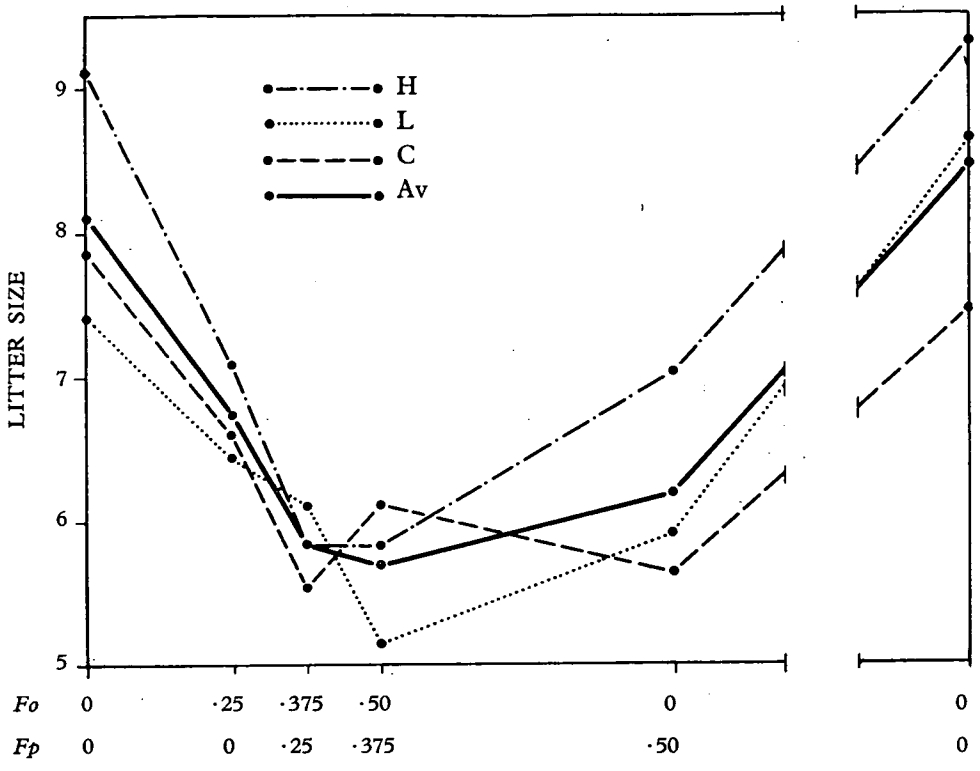


Fig. 3. Litter size plotted against inbreeding coefficient of the offspring (F_o). Inbreeding coefficient of parents (F_p) also shown. Group H—previously selected for high litter size; Group L—previously selected for low litter size; Group C—not previously selected; Av—average of all groups.

estimate of the initial effect of crossbreeding. Had the inbred parents borne inbred rather than crossbred young, their estimated litter size would be somewhere in the region of 5·0, assuming a linear decline. This would indicate that the real effect of crossbreeding in the litter was to increase litter size by rather more than one mouse. Nevertheless, this increase was considerably smaller than expected if we were to extrapolate from the results of the first inbred generation, where a bigger change in litter size occurred for only half the change in inbreeding coefficient. The anomaly does not end here. In the second inbred generation, the additional effect of inbreeding in the dam was barely perceptible over the expected effect of further inbreeding in the young. Yet, in generation X_2, the effect of crossbreeding in the dam was to

increase litter size by 2·27 over the previous generation. The data suggest that inbreeding may impose a limit on the dam's potential fertility, and that no amount of heterozygosity in the young would increase litter size above a certain level. In outbred dams, on the other hand, any reduced viability through inbreeding in the unborn young would be fully revealed in the reduced litter size at birth.

The mean litter size of the crossbred mice in generation X_2 is 0·35 of a mouse higher than in the original outbreds, generation O. The comparison of these two means is of prime importance and represents one major interest of the experiment. The difference is not significant at the 5% level, despite the slight involuntary between-line selection during inbreeding, mentioned earlier. Over the period of the experiment, generation means of the outbred control varied between 7·00 and 8·17, which makes a difference of 0·35 appear unimportant. There is therefore no reason to suppose that natural selection operating within lines during inbreeding has had any effect on the mean performance of the derived crossbred population.

It is perhaps of some interest to consider separately the three groups of mice from which the inbred lines were derived. These groups, though initially of common origin, had become differentiated through selection for high and low litter size, the third group being an unselected control. It is conceivable therefore that the effect of inbreeding on litter size could well be different in the different groups. Considering first the two groups that had previously been selected, it seems that after three full-sib matings, litter size had declined in both by approximately the same proportionate amount to about two-thirds of the initial litter size. But the increase on crossing was relatively greater in the group erstwhile selected for small litters. However, the standard errors of all these estimates of group means were of the order of 0·4 of a mouse, and without any more elaborate statistical analysis it is clear that apparent differential trends of the magnitude observed would not be significant. In the group that had not previously been selected, litter size increased during the last generation of inbreeding and fell again when the lines were crossed. This does not accord with expectation nor with the behaviour of the other two groups. It seems probable that the estimate of the mean of the I_3 generation in this group is spuriously high either through sampling errors or through some short-term environmental influence which the other groups did not encounter.

Litter size, as mentioned earlier, is markedly affected by the weight of the dam, and the picture is therefore not complete without the examination of this correlated character for possible changes during the experiment. If weight were to decrease on inbreeding with a subsequent increase on crossing, this would have obvious repercussions on the interpretation of the observed effects on litter size. Because of this possibility, the weight of the females was recorded at 6 weeks, the approximate age at mating. The mean weight is shown in Table 1. The first conclusion is that 6-week weight did not change in any systematic manner with changes in heterozygosity. Secondly, such changes as were observed were so small that any correction of litter size for dam's weight would only have a trivial effect. Though body weight in standardized litters is known to decline on inbreeding, it seems that in this experiment the depression was more or less balanced by the advantage gained

through a simultaneous reduction in litter size. Likewise, when the lines were crossed, the potential increase in body weight was nullified by the increase in litter size.

(*b*) *Differentiation between inbred lines*

The classical theory of inbreeding indicates that inbred lines become differentiated, with a corresponding increase in uniformity within lines. The mathematical expressions for the variances between and within lines are $2F\sigma_A^2$ and $(1-F)\sigma_A^2$, respectively, where σ_A^2 is the additive genetic variance in the initial population and F is Wright's coefficient of inbreeding. At complete inbreeding the initial genetic variation is thereby doubled, and it all appears between lines. However, these expressions are true only if all the variance is additive; they will not hold where dominance and epistatic deviations exist, and in most instances the observed result on inbreeding will differ from expectation based on an additive model.

The theoretical treatment of the effect of inbreeding on variation in a non-additive situation has not been developed fully, but Robertson (1952) has examined the consequences on variation due to rare recessive genes. He showed that the within-line variance in such a case would increase on inbreeding until F is in the region of 0·5 and then decline. The between-line variance will also increase, but only slowly at first as the increase is proportional to F^3. Robertson shows further that the same general conclusions will probably apply to genes showing over-dominance.

It appears therefore that, in an unknown genetic situation, changes in within-line and between-line variances are unpredictable, and for this reason every empirical observation is of some value. The results obtained from the present work are summarized in Table 2. The data from generation X_1, where the offspring are crossbred, are not included as the variance observed cannot be partitioned in a simple manner into within-line and between-line components.

Table 2. *Variance components within and between inbred lines*

	Generation		
	I_1	I_2	I_3
Within-line component	5·19	6·70	3·16
Between-line component	1·49	0·08	2·81

With only three points available for examination, it is clearly impossible to establish any definite trend. Further, as the estimates of the within-line and between-line components are necessarily negatively correlated, it becomes difficult to deduce anything about their interrelationship. The values obtained for the I_2 generation must be spurious, for on no model would the differentiation between lines vanish so suddenly only to re-emerge in the subsequent generation. But if any reliance can be placed on the other estimates, it seems that the total variation is being repartitioned in the direction of increasing the differentiation between lines.

(c) *Analysis of variance in crosses*

The data have to be analysed in two distinct classifications. The first of these concerns crosses (irrespective of whether the cross is AB or BA), reciprocal members of the same cross, and error variance. In the second classification, the variance is partitioned between dam-lines, between sire-lines and the interaction between them. We shall consider the two classifications in this order.

The error variance is of course common to both. This was estimated from twenty-two duplicate crosses in the first cross, and forty triplicate matings in the second. In these replicate crosses, parents of the same sex were always taken from the same line. It proved to be immaterial whether replicates were taken from the same litter or from different litters from the same cross. Both analyses were made within the three major groups that constituted the experimental population. 'Group' refers to a set of lines of common origin. The results of the first analysis are shown in Table 3. There seems to be little evidence of variation between crosses in either generation. This indicates that no effective selection of good crosses from the array of possible ones could be made.

Table 3. *Analysis of variance in crosses*

	First cross			Second cross		
	d.f.	M.S.	*P*	d.f.	M.S.	*P*
Total	106	7·18		146	5·03	
Between groups	2	19·46	$> 0{\cdot}05 < 0{\cdot}10$	2	42·70	$< 0{\cdot}001$
Within groups	104	6·94		144	4·51	
Between crosses	44	7·08	$> 0{\cdot}20$	38	5·91	$> 0{\cdot}50$
Within crosses	60	6·84		106	4·01	
Between reciprocals	38	5·93	$> 0{\cdot}20$	26	6·54	$< 0{\cdot}01$
Between replicates	22	8·41		80	3·18	

The influence of maternal effects on litter size is illustrated by the significant difference between reciprocals in the second cross. No such difference could be established in the first cross, probably because of the magnitude of the error mean square. The large error variance, especially in the first cross, is a disconcerting feature of the data. This suggests that no precise estimates of the components of variance involved could be obtained, without large-scale experimentation.

The second analysis attempted to partition the variance between lines, used both as male and female parents, and to measure the interaction between them. This should enable us to distinguish between the 'general combining ability' of a line, which can be defined as the average performance of crosses between that line and all other lines, and the 'special combining ability' of a cross, measured by the deviation of the performance of that cross from the expectation based on the general combining abilities of its parent lines. The variation in the general combining ability of lines will be represented by the sum of two components of variance, that

between dam-lines and the one between sire-lines. The variation in specific combining ability will be the interaction component of variance.

The method whereby the components were estimated was somewhat complicated, owing to the non-orthogonality of the system of crossing and also because a dam-line was crossed only to some of the sire-lines, and vice versa. The analysis is therefore not presented in any detail, but the principle involved is explained by Henderson (1953). The estimates obtained for the components in the two generations of crossing are shown in Table 4.

Table 4. *Components of variance of litter size in crosses*

Component	First cross	Second cross
Between sire-lines	0	1·08
Between dam-lines	0	0·36
Interaction	0	0·07
Error	8·41	3·18

In the first cross, all the components except error took a small negative value, giving zero as the best estimate in each case. In the second cross, the interaction component was very small indeed, indicating that, in this particular situation, specific combining ability is practically non-existent, and certainly very small compared with the general combining abilities of the lines. Because of their composition, the appropriate mean squares could not be adequately tested for significance level.

The order of magnitude of these components compared to the error variance again indicates that for accurate estimation the scale of the experiment is inadequate. But even after allowing for large error variance, there seems to be little evidence of any useful variation between crosses, indicating that selection between crosses would be ineffective.

5. DISCUSSION

The interpretation of the experimental data has been rendered somewhat imprecise by the complexity of the character of litter size. The difficulties involved can be attributed in no small measure to the dual genetic determination of the character, as the relative contributions of the dam and of the litter itself are seldom clearly distinguishable. In addition, we have strong maternal effects on litter size, and their interplay with inbreeding depression adds further intricacies. The examination of the underlying genetic situation will therefore be severely limited in its scope until such time as the constituent factors of litter size are more perfectly understood.

To some extent we have seen the genotype of the dam and the genotype of the litter acting on litter size separately. At the commencement of inbreeding, reduced viability of the unborn litter had a marked effect which was only partly recovered when the lines were crossed. By then, crossbreeding in the dam appeared to be of

predominating importance in increasing litter size, but the effect of inbreeding in the dam, when first introduced, had been barely perceptible. The explanation may be, in part at least, a maternal effect restricting litter size in inbred mothers irrespective of the heterozygosity of the young. The elimination of lethals with a heterozygote advantage cannot be invoked, for untimately the mean litter size of the original outbred population was restored when the crossbred mice were used as parents. In the absence of selection, this is what theoretical considerations lead us to expect, for unselected inbred lines could then be regarded as a sample of the gametes of the original outbred population. A random cross would therefore represent one individual of the original outbreds.

It is only fair to admit that the apparent contradiction mentioned above could have arisen if the mean litter size for either the I_2 or X_1 generation had been spuriously low. Yet, this seems unlikely, for other workers report analogous findings. Eaton (1953) noted when he crossed inbred lines of mice that the effect on litter size of crossbreeding in the dam was much larger than the effect of crossbreeding in the litter, if inbreeding had proceeded for less than six generations. The assessment of the other point, namely the I_2 generation, is confirmed by Bowman and Falconer (1960), who with the same stock of mice in the same laboratory found a similar rate of decline on inbreeding. When all these complementary phenomena are considered together, the possibility of sampling error becomes reduced, and it would seem that the decline in litter size is not linearly related to inbreeding when its effect in the dam and in the litter are considered separately.

The effect on crossbred performance of any natural selection operating within lines during the inbreeding stage appears now to be unimportant. The improvement in fertility normally associated with crossing subsequent to inbreeding must therefore be ascribed to some other form of selection, as a result of which many of the poorer genotypes would not be represented in the crossbred population. Hybrids between a random array of inbred lines have no intrinsic merit except to the extent that the population was selected during inbreeding. For certain characters, inbreeding and crossing may well provide means of rapid selection, whether natural or artificial, that might not otherwise be possible. Apart from this possibility, the only advantage of the system would be the ability to replicate any desired cross at will.

The lack of variation between the means of the crosses was somewhat unexpected, for the inbred lines were clearly differentiated in the last generation of inbreeding. The probable explanation lies in the use of partly inbred material. It may be shown from a paper by Robertson (1952) that the expected variance between the means of line crosses is $F\sigma_A^2 + F^2\sigma_D^2$, where F is Wright's coefficient of inbreeding when the lines are crossed, σ_A^2 is the additive component of variance, and σ_D^2 is the variance due to dominance. Hence, in this particular experiment only half the additive and a quarter of the dominance variance was available. The additive genetic component of variance in this stock of mice is of the order of 1·5. No similar estimate can be made of the variance due to dominance, but only in special circumstances would it be much greater than the additive component (see, for instance, Mather, 1949). It

is apparent therefore that compared with the error variance observed, these estimates of the genetic sources of variation to be expected, when divided between two generations, become very small. This indicates that before any useful selection could be made between crosses, not only should the experiment be on a larger scale, but also the level of inbreeding should be advanced well beyond 50 per cent.

It can also be shown, from Robertson's paper, that the term $F\sigma_A^2$ represents the component of variation due to the general combining ability of the lines, while $F^2\sigma_D^2$ is a component ascribable to special combining ability of lines in particular crosses. It can therefore be seen that until the level of inbreeding is well advanced, special combining ability will always play a subsidiary role to the general combining ability of the lines, unless the dominance variance is exceptionally large compared to the additive genetic component. Such a situation might occur if overdominant loci, with genes at intermediate frequencies, were contributing largely to the total variance. Employing a somewhat subjective assessment and applying the law of parsimony, it seems that overdominance at a number of loci was not encountered in this study.

The application of these results will be limited to situations of similar genetic control, but in conclusion, inbreeding and crossing as a method of improving a character such as the one described in this paper will not prove useful unless lines at a fairly advanced level of inbreeding are maintained. Even then, many if not most crosses may not be successful in increasing litter size. In view of this, it is encouraging that the within-family selection experiment, carried out on the same stock in this laboratory, has by now produced a substantial difference in litter size between the high and low lines (Falconer, 1955 and unpublished). It has just been shown that there is no reason to suppose that the character is controlled by many overdominant loci, which would preclude the successful outcome of a selection programme. Only in such circumstances would inbreeding and crossing be a better method of improving the character.

SUMMARY

1. The experiment was designed to provide basic information relevant to the utilization of heterosis in animal improvement. The character studied was the size of the first litter in mice.

2. Thirty inbred lines were crossed at random when the inbreeding coefficient reached 0·50 (three full-sib matings). The lines had been inbred without selection except for natural selection operating with lines.

3. The mean litter size of the crossbred mice did not exceed that of the outbred population from which the inbred lines had been derived. This indicates that the increased litter size normally associated with crossbred mice must be ascribed to some form of selection other than within-line natural selection.

4. Estimates were obtained of the variance components associated with general and special combining abilities. As anticipated, these estimates were very small, especially those relating to special combining ability. Before selection between crosses becomes possible, high levels of inbreeding must be achieved.

This work was carried out while on leave of absence from the Agricultural Research Council Animal Breeding Research Organization, and I am indebted to the Director of the Organization, Dr H. P. Donald, for his support.

I am grateful to Dr D. S. Falconer for his advise throughout the course of this work, to Dr A. Robertson and Dr B. Woolf for statistical advise, and to Dr E. C. R. Reeve for helpful comments on the manuscript.

REFERENCES

BOWMAN, J. C. and FALCONER, D. S. (1960). Inbreeding depression and heterosis of litter size in mice. *Genet. Res.* **1**, 262–274.

DONALD, H. P. (1955). Controlled heterozygosity in livestock. *Proc. Roy. Soc.* B, **144**, No. 915, 192–203.

EATON, O. N. (1953). Heterosis in the performance of mice. *Genetics*, **38**, 609–629.

FALCONER, D. S. (1955). Patterns of response in selection experiments with mice. *Cold Spring Harbor Symp. Quant. Biol.* **20**, 178–196.

HENDERSON, C. R. (1953). Estimation of variance and covariance components. *Biometrics*, **9**, 226–252.

MATHER, K. (1949). *Biometrical Genetics*. London: Methuen.

ROBERTSON, A. (1952). The effect of inbreeding on the variation due to recessive genes. *Genetics*, **37**, 189–207.

PAPER 2

Effect of inbreeding on ovulation rate and foetal mortality in mice

Genet. Res. 1, 422 - 430. 1960.

by

D.S. FALCONER & R.C. ROBERTS.

Genet. Res., Camb. (1960), **1**, *pp.* 422–430
Printed in Great Britain

Effect of inbreeding on ovulation rate and foetal mortality in mice

By D. S. FALCONER and R. C. ROBERTS

Agricultural Research Council Unit of Animal Genetics, Institute of Animal Genetics, Edinburgh, 9

(*Received* 18 *February* 1960)

INTRODUCTION

The reduction of the litter size of mice resulting from inbreeding was described in two earlier papers (Roberts, 1960; Bowman & Falconer, 1960). The number of live young born in first litters was taken as a measure of litter size, and this was found to decline in a very regular manner as inbreeding increased: at an inbreeding coefficient of 50% the reduction was about 2½ young per litter. Crosses between lines inbred to 50% showed that the reduction was attributable in part to inbreeding in the litters and in part to inbreeding in the mother bearing the litter. The present paper is concerned with the developmental stage at which the reduction of litter size takes place. The reduction could arise from a reduced ovulation rate, an increased loss of eggs or embryos before implantation, or an increased mortality of embryos at any stage between implantation and birth. Dissections of pregnant females were made in order to identify the stage at which the losses from inbreeding take place. The females were dissected 16 days after insemination, and counts were made of (1) the number of corpora lutea, as a measure of the number of eggs ovulated, (2) the total number of implantation sites, and (3) the number of embryos alive at 16 days of gestation. The difference between (1) and (2) indicated pre-implantation losses and the difference between (2) and (3) indicated post-implantation mortality. Further, a comparison of the number of embryos alive at 16 days in the dissected females with the number of live young born to comparable females, allowed to bring their litters to term, provided a measure of perinatal loss, i.e. losses occurring in the last 2 or 3 days of gestation and between the birth and the recording of the litter. Our interest was primarily in the maternal contribution to the inbreeding depression of litter size, i.e. in the stage at which inbreeding of the mother reduces the size of the litter she bears. For this reason most of the comparisons made were between inbred and non-inbred parents, both with non-inbred embryos.

STOCKS AND TECHNIQUES

Three series of dissections were made, all on females of the same basic stock. The dissections and counts of the first series were made by Roberts, those of the second and third series by Falconer. The details of the three series were as follows:

Series I. The mice came from the inbred lines described by Roberts (1960).

Twenty-one independently inbred lines were represented among the females dissected. Five groups of females were arranged according to whether the mother, the father, or the embryos were inbred or non-inbred. The inbreeding coefficients of the mice used in the five groups are given with the results in Table 1. The inbreeding coefficients were either 50% or 59%, or zero.

Series II. The mice came from an experiment on selection for sex-ratio (Falconer, 1954). Nine independently inbred lines were represented among the females dissected. The mice for dissection were 50% inbred and they were mated to males similarly inbred but of a different line. The embryos were thus non-inbred. Comparisons were made with a contemporary group of non-inbred females, produced from line-crosses and mated to similar males from a different line-cross. Thus the only difference between the two groups compared in the series-II dissections was in the inbreeding of the parents.

Series III. The inbred mice were derived from six of the independently inbred lines described by Bowman & Falconer (1960). The females were mated to males from a different line, in the same way as in Series II, so that only the parents and not the embryos were inbred. The coefficient of inbreeding of the dissected females was 63%. The non-inbred mice for comparison came from three sources: a line selected for large litters, a line selected for small litters, and a control line maintained without selection. All three lines had been maintained with minimal inbreeding, and were in the fourth generation of the selection experiment.

All the dissections were made 16 days after insemination, indicated by the presence of a vaginal plug. The uteri were opened and the numbers of implantation sites and of live embryos were noted. Few embryos were found that had died at an identifiable stage of development. For this reason the different stages of post-implantation death were not distinguished in the analysis of the results. The numbers of corpora lutea were counted by examination of the ovaries under a low-power binocular microscope. Exact correspondence between the number of corpora lutea counted and the number of eggs shed was not to be expected, because it was difficult, particularly when the corpora lutea were numerous, to distinguish between one large corpus luteum and two adjacent and partially confluent ones. On the whole the number counted is probably an underestimate rather than an overestimate of the ovulation rate. The corpora lutea counted in some mice of the first series of dissections were compared with egg counts made on a comparable group of females by Dr A. W. H. Braden. The mean number of corpora lutea among 38 females was $10{\cdot}6 \pm 0{\cdot}27$ and the mean number of eggs among 32 females was $10{\cdot}1 \pm 0{\cdot}29$. The close correspondence between the two suggests that, in this series at least, the corpora lutea counts gave a good estimate of the ovulation rate. Among the non-inbred mice of series II and III, 9 out of 74 pregnancies (12·2%) showed an excess of implants over corpora lutea (see Table 3). This, however, is an underestimate of the percentage of errors because about 60% of pregnancies showed some loss of eggs before implantation, and an error of counting would only be revealed when there was no loss. As a very rough estimate we may say that probably about 30% of the counts were too low.

Since, however, we are to compare groups of mice counted in the same way, the bias introduced should not seriously affect the conclusions to be drawn.

There were a few mice with corpora lutea but no implantations. These were excluded, on the grounds that a total loss of eggs would be followed by another ovulation before the dissections were made, so the corpora lutea counted would not represent the eggs that had been lost.

RESULTS

The mean numbers of corpora lutea, implants, and live embryos are given in Table 1. Inspection of these figures shows clearly that inbreeding does not reduce the number of corpora lutea but it does reduce the numbers of implants and of live embryos. Let us, however, examine the results more closely. The five groups of females dissected in Series I are shown separately at the top of Table I. None

Table 1. *Mean numbers of corpora lutea, implants, and live embryos at* 16 *days' gestation* (± *standard errors*). *N is the number of females dissected*

Group	Inbreeding coefficients (%) ♀♀	♂♂	Embryos	N	Corpora lutea	Implants	Live embryos
Series I							
A	50	50	59	27	10·0 ± 0·34	8·4 ± 0·48	7·1 ± 0·48
B	50	50	0	29	9·9 ± 0·32	7·6 ± 0·54	6·7 ± 0·52
C	59	0	0	30	10·2 ± 0·51	8·8 ± 0·50	7·8 ± 0·53
D	0	59	0	28	9·9 ± 0·26	8·9 ± 0·37	8·1 ± 0·52
E	0	0	0	30	10·3 ± 0·33	9·1 ± 0·40	7·8 ± 0·54
A+B+C	50–59	—	—	86	10·0 ± 0·23	8·3 ± 0·29	7·2 ± 0·30
D+E	0	—	0	58	10·1 ± 0·21	9·0 ± 0·27	7·9 ± 0·37
				t =	*0·22*	*1·79*	*1·57*
				P =	*0·9*	*0·1*	*0·2*
Series II	50	50	0	13	10·9 ± 0·58	8·5 ± 0·69	7·3 ± 0·74
	0	0	0	15	11·7 ± 0·46	11·3 ± 0·57	9·8 ± 0·66
				t =	*1·09*	*3·15*	*2·54*
				P =	*0·3*	*0·01*	*0·05*
Series III	63	63	0	17	12·5 ± 0·82	7·9 ± 0·82	6·4 ± 0·80
(H)*	0	0	0	23	10·0 ± 0·41	9·4 ± 0·28	8·2 ± 0·41
(L)*	0	0	0	18	11·9 ± 0·80	9·9 ± 0·71	8·9 ± 0·69
(C)*	0	0	0	18	9·1 ± 0·29	8·0 ± 0·42	7·4 ± 0·47

* H, L and C are the lines selected for high and low litter size and unselected control, respectively.

of the differences between the groups, considered separately, are clear enough to allow us to draw firm conclusions about the different effects of inbreeding in the mother, the father, or the embryos. In particular, the only comparison that contains information about the effect of inbreeding in the embryos (groups A

and B) shows no reduction of the numbers of implants or embryos. The apparent absence here of any effect of inbreeding in the embryos cannot, however, be given much weight because the numbers are not very large and because an increase of the numbers born from crosses between highly inbred lines has often been observed (see, for example, Eaton, 1953). Our chief interest here is in the effect of inbreeding in the mother of the litter, and for this purpose we may combine the groups according to whether the mother was inbred or non-inbred. This gives the same type of comparison as is made in the dissections of Series II and III. The combined groups are given also in Table 1. The numbers of corpora lutea are now almost identical, showing that the ovulation rate is not affected by inbreeding. The numbers of implants and of live embryos are both lower in the inbred females than in the non-inbred, though neither of these differences reach a fully convincing level of significance.

The results of the Series II dissections are quite clear. There is a small but non-significant difference in corpora lutea. Both the implants and the live embryos are fewer in the inbred than in the non-inbred females, and both differences are significant at the 5% level. In the Series III dissections the inbred females have more corpora lutea, but again fewer implants and live embryos. The three non-inbred groups, however, differ significantly between themselves and therefore cannot be combined. These differences were associated with differences of body weight.

The significance tests of the differences between group-means given in Table 1 and also in Table 2 followed the method given by Snedecor (1956, pp. 97–98) which does not assume equality of variance. In fact, the variances of the inbred and non-inbred groups differed significantly in almost every comparison, the inbreds being the more variable. It is not possible, however, to draw genetical conclusions from this fact because the expected changes of the genetic variance at intermediate levels of inbreeding cannot be predicted unless all the variance is additive (Robertson, 1952). For this reason the variances will not be further discussed.

From the three series of dissections we may conclude at this stage that inbreeding does not reduce the ovulation rate as measured by the number of corpora lutea. The reduction of litter size at birth must therefore arise from losses of eggs or embryos either before or after implantation.

Losses

Consideration of the losses of eggs or embryos makes the picture clearer. The pre- and post-implantation losses in the three series of dissections are given in Table 2. The dissections of Series I are here grouped according to the inbreeding of the female parent. The three groups of non-inbred females in Series III are here combined because the differences between them, though still just significant, are much less. All three series agree in showing a greater pre-implantation loss in inbreds than in non-inbreds, but no difference in the post-implantation losses. The significance of the difference of pre-implantation loss does not quite reach the 5% level in Series I, but in Series II and III it reaches the 1% level. There can be

Table 2. *Mean numbers of pre- and post-implantation losses (± standard errors). The percentage losses show respectively the percentage of corpora lutea not represented by implantation sites and the percentage of implants not represented by live embryos at 16 days*

Series	Group	N	Loss: Pre-implantation		Loss: Post-implantation	
I	A+B+C (♀♀ inbred)	86	1·77 ± 0·245	17·6%	1·07 ± 0·179	12·9%
	D+E (♀♀ crossbred)	58	1·12 ± 0·219	11·1%	1·04 ± 0·159	11·5%
			t = 1·99			
			P = 0·05			
II	Parents inbred	13	2·46 ± 0·765	22·5%	1·15 ± 0·355	13·6%
	Parents outbred	15	0·47 ± 0·336	4·0%	1·47 ± 0·390	13·0%
			t = 2·39			
			P = 0·05			
III	Parents inbred	17	4·65 ± 1·010	37·1%	1·47 ± 0·355	18·7%
	Parents outbred	59	1·24 ± 0·237	12 0%	0·93 ± 0·159	10·2%
			t = 3·29			
			P = 0·01			

little doubt that the difference is real in all three series. The conclusion is, therefore, that inbreeding in the female parent increases the loss of eggs or early embryos between ovulation and implantation, but it does not increase the mortality of embryos after implantation.

If the groups of Series I are arranged according to the inbreeding of the male parent, the pre-implantation losses are a little greater with inbred than with non-inbred fathers (1·70 against 1·28), but the difference is insignificant. The pre-implantation loss from inbred mothers is therefore more probably attributable to the females themselves than to a failure of fertilization caused by inbreeding of the male parent.

The distributions of pre-implantation losses are given in Table 3. The higher mean loss in inbreds is due to a larger number of large losses rather than to a difference in the modal loss: in other words, to a drawing out of the upper tail of

Table 3. *Distributions of pre-implantation losses, i.e. numbers of dissections exhibiting each degree of loss. 'Negative' losses represent deficiencies of corpora lutea due to miscounting*

	Number of eggs lost (= excess of corpora lutea over implants)															
	−2	−1	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Series I																
Inbred ♀♀		1	30	18	14	12	5	0	1	0	1	3	1			
Outbred ♀♀	1	1	24	16	7	4	1	3	0	1						
Series II and III																
Inbred ♀♀			6	7	2	2	3	2	1	1	2	1	1	1	0	1
Outbred ♀♀	1	8	23	23	7	5	4	0	0	2	1					

the distribution. The skewed distributions make the exact level of significance of the *t*-tests a little uncertain, but the reality of the difference between inbred and non-inbred females can hardly be questioned.

Nothing has yet been said about the perinatal losses. Not all the groups of dissected females had suitable females available for comparison of the number of young born alive. Those available for valid comparison are shown in Table 4.

Table 4. *Comparison of live embryos at* 16 *days and live young at birth in litters of comparable females*

		Live embryos		Live births	
Series	Parents	N	Mean ± s.e.	N	Mean ± s.e.
I B	Inbred	29	6·7 ± 0·52	107	6·2 ± 0·067
E	Outbred	30	7·8 ± 0·54	147	8·5 ± 0·034
II	Inbred	13	7·3 ± 0·74	21	7·4 ± 0·70
	Outbred	15	9·8 ± 0·66	27	8·3 ± 0·38

The evidence, as far as it goes, points to the perinatal losses being very small, and substantially the same in inbred as in non-inbred mothers.

Connexion between ovulation rate and body weight

The connexion between ovulation rate and body weight has an important bearing on the interpretation of the fact that ovulation rate did not decline with inbreeding. This will be explained below, in the Discussion, but the data are given here. The influence of body weight on the number of corpora lutea was examined in the dissections of Series I. The females were weighed at 6 weeks of age and were mated soon after, so that the ovulation corresponding to the corpora lutea counts took place between the ages of about 7 and 9 weeks. The regression of corpora lutea on 6-week weight was 0·244 ± 0·063 corpora lutea per gram. (The regressions did not differ significantly between the groups and this is the pooled value.) There is, thus, a positive association between body weight and ovulation rate. Comparisons of ovulation rate between inbred and outbred females should therefore take into account any differences of body weight. The mean weights of the females in the three series of dissections are given in Table 5, together with the mean ages at ovulation and the numbers of corpora lutea, from Table 1, to facilitate comparison. It will be seen that the weights of the inbred and outbred females of series I were very nearly the same, and so were the ovulation rates. The inbred females of Series II were 2 g. lighter than the outbred females at the time of ovulation (the weights were not recorded at 6 weeks) and the ovulation rate was lower. In Series III, as in Series I, the mean 6-week weight of the inbred females was not less than that of the outbred females. The differences of weight between the groups were, however, rather larger than in Series I, and the inbred females were substantially older than the outbred ones at the time of ovulation. The comparisons of ovulation rates in this series are therefore less reliable

Table 5. *Weights and ages of the females dissected*

Series	Group	Corpora lutea	Weight at 6 weeks (g.)	Weight at insemination (g.)	Age at insemination (days)
I	Inbred ♀♀ (A+B)	10·0	21·7	—	53
	(C)	10·2	21·7	—	59
	Outbred ♀♀ (D+E)	10·1	21·4	—	58
II	Inbred ♀♀	10·9	—	27·1	93
	Outbred ♀♀	11·7	—	29·0	95
III	Inbred ♀♀	12·5	22·9	—	103
	Outbred ♀♀ (H)	10·0	20·6	—	59
	(L)	11·9	23·8	—	69
	(C)	9·1	22·0	—	60

DISCUSSION

The fact that the ovulation rate, measured by the number of corpora lutea, was not affected by inbreeding calls for some comment. It would be reasonable to expect that body weight would decline on inbreeding, and, since ovulation rate is positively correlated with body weight, a reduction of ovulation rate on inbreeding might reasonably be expected. In fact, however, the weights of the mice used in this work, except those of Series II, did not decline on inbreeding for the following reason (see Roberts, 1960). The weights of mice, both at weaning and subsequent ly, are inversely correlated with the number reared in the litter, so that mice reared in small litters are larger than mice reared in large litters. (See, for example, Falconer, 1955.) No adjustment of the litter size at birth was made during the inbreeding of the mice dissected in Series I and III, the mothers being left to suckle all the mice to which they gave birth. But the numbers born declined as inbreeding proceded, and consequently the weights tended to increase. Presumably there was a contrary tendency for the weights to decrease from reduced pre-weaning nutrition and slower growth. These two opposing tendencies counterbalanced each other and the weights remained substantially constant. By this fortunate coincidence we were able to assess the effects of inbreeding on the ovulation rate without the disturbing influence of associated changes of body weight. The mice dissected in Series II, however, were differently treated. Litters were standardized at birth to four young during the inbreeding, and the compensatory effect of declining litter size was here absent. The inbred females were lighter than the outbred females and their ovulation rate was lower, though neither difference was statistically significant.

Thus the conclusion that inbreeding does not influence ovulation rate requires qualification because it refers only to the rather special circumstances where the effect of inbreeding on growth is not apparent. Under other circumstances, which allow body size to decline with inbreeding, the ovulation rate would almost certainly decline too. The conclusion, translated into genetic terms, is that genes that

influence ovulation rate without affecting body size do not show directional dominance, but genes that influence ovulation rate as a consequence of their effect on body size may show directional dominance.

It is interesting to note that the effect of inbreeding on the ovulation rate of pigs appears to be different. The ovulation rate is reduced by inbreeding (Squiers *et al.*, 1952; King & Young, 1957), and this—rather than pre- or post-implantation loss—is the chief cause of the reduction of litter size. One cannot invoke a reduction of body size as an explanation because the ovulation rate was not found to be significantly correlated with body weight (King & Young, 1957). An explanation of the difference between pigs and mice might perhaps lie in their past histories of selection. The directional dominance exhibited by the genes affecting the ovulation rate of pigs may be the consequence of selective pressure in the past directed toward increased litter size; there has probably been much less selection of this sort in the past history of laboratory mice.

Reverting now to mice: the conclusion reached from the work described here is that the reduction of litter size in inbred mothers is mainly due to an increased loss of eggs or embryos before implantation. This loss could be caused by failure of fertilization or by failure of the fertilized eggs to implant. Since the inbreeding of the male did not influence the pre-implantation loss, any failure of fertilization would have to be attributed to the female. Losses attributable to the female might arise from (i) the production of abnormal eggs, (ii) impaired transport of the sperm to the site of fertilization, or (iii) failure of implantation in consequence of an impairment of endocrine function. The first seems unlikely because Braden (1957) found the incidence of abnormal eggs was no higher in inbred than in non-inbred females. The second is possible because only quite a small number of sperm reach the site of fertilization (Braden, 1958) and it would not be unreasonable to suppose that impaired transport might reduce their number enough to leave some eggs unfertilized. Nevertheless, failure of the fertilized eggs to implant, in consequence of an impaired endocrine function, seems to be the more likely cause of the increased pre-implantation loss in inbred females.

SUMMARY

Dissections were made of 16-day-pregnant female mice with the object of discovering the developmental stage at which litter size is reduced by inbreeding. Counts were made of the numbers of corpora lutea, implantation sites, and live embryos, and comparisons were made between females with inbreeding coefficients of 50–60% and non-inbred females. Except in one group the embryos were all non-inbred, so that the comparisons showed the effect of inbreeding in the mother of the litter. No influence of inbreeding in the male parent was found.

The only difference found between inbred and non-inbred females was in the number of eggs or embryos lost before implantation. The greater pre-implantation loss in inbred females was enough to account for the smaller number of young born alive in their litters.

There was no difference between the inbred and non-inbred females in the ovulation rate, measured by the number of corpora lutea, or in the post-implantation mortality of the embryos.

There was a positive correlation between ovulation rate and weight at 6 weeks. For reasons explained in the Discussion, the inbred females did not differ in weight from the non-inbred females. If, under other conditions, the weight declined on inbreeding, the ovulation rate would be expected to decline also.

REFERENCES

BOWMAN, J. C. & FALCONER, D. S. (1960). Inbreeding depression and heterosis of litter size in mice. *Genet. Res.* **1**, 262–274.

BRADEN, A. W. H. (1957). Variation between strains in the incidence of various abnormalities of egg maturation and fertilization in the mouse. *J. Genet.* **55**, 476–486.

BRADEN, A. W. H. (1958). Variation between strains of mice in phenomena associated with sperm penetration and fertilization. *J. Genet.* **56**, 1–11.

EATON, O. N. (1953). Heterosis in the performance of mice. *Genetics*, **38**, 609–629.

FALCONER, D. S. (1954). Selection for sex ratio in mice and *Drosophila*. *Amer. Nat.* **88**, 385–397.

FALCONER, D. S. (1955). Patterns of response in selection experiments with mice. *Cold Spr. Harb. Symp. quant. Biol.* **20**, 178–196.

KING, J. W. B. & YOUNG, G. B. (1957). Maternal influences on litter size in pigs. *J. agric. Sci.* **48**, 457–463.

ROBERTS, R. C. (1960). The effect on litter size of crossing lines of mice inbred without selection. *Genet. Res.* **1**, 239–252.

ROBERTSON, A. (1952). The effect of inbreeding on the variation due to recessive genes. *Genetics*, **37**, 189–207.

SNEDECOR, G. W. (1956). *Statistical Methods* (5th ed.). Iowa State College Press, Ames, Iowa.

SQUIERS, C. D., DICKERSON, G. E. & MAYER, D. T. (1952). Influence of inbreeding, age and growth rate of sows on sexual maturity, rate of ovulation, fertilization and embryonic survival. *Res. Bull. Mo. agric. Exp. Sta. No.* 494.

PAPER 3

Analysis of differences in the number of eggs shed by the two ovaries of mice during natural oestrus or after superovulation.

J. Reprod. & Fertil. 2, 418-437. 1961.

by

D.S. FALCONER, R.G. EDWARDS, R.E. FOWLER & R.C. ROBERTS.

Journal of Reproduction and Fertility (1961) 2, 418-437

ANALYSIS OF DIFFERENCES IN THE NUMBERS OF EGGS SHED BY THE TWO OVARIES OF MICE DURING NATURAL OESTRUS OR AFTER SUPEROVULATION

D. S. FALCONER*
Institute of Animal Genetics, Edinburgh

R. G. EDWARDS
National Institute for Medical Research, London, N.W.7

R. E. FOWLER AND R. C. ROBERTS*
Institute of Animal Genetics, Edinburgh

(*Received* 12*th April* 1961)

Summary. The variation in the number of eggs shed by the two ovaries of mice has been examined by statistical analyses of 697 egg counts and 390 corpora lutea counts, made on mice from a variety of outbred strains, both after natural oestrus and after oestrus induced in adults by pregnant mares' serum (PMS) and human chorionic gonadotrophin (HCG). The numbers of eggs or corpora lutea were distributed between sides approximately at random, the variation conforming fairly closely to a binomial distribution. This was true even after superovulation. There was, however, a slight but significant excess of variation between sides over the random amount in the egg counts, particularly after natural ovulation. Corpora lutea counts differed from egg counts in showing a slight but significant reduction of the variation below the random amount. Several possible reasons for these small deviations from a random distribution are discussed.

The correlation between the numbers of eggs shed by the two ovaries was negative after natural ovulation but positive after superovulation. This difference can be fully accounted for by the random distribution between sides together with the differences of mean and variance between natural ovulation and superovulation. The variation of total egg number was proportional to the mean egg number after natural ovulation. The variation after superovulation was much higher than after natural ovulation, even when the difference of mean was taken into account, and the greater variation of total egg number caused the correlation between sides to be positive after superovulation.

INTRODUCTION

It is a curious, though familiar, fact that the numbers of eggs shed by left and right ovaries of naturally ovulating mice are negatively correlated. That is to

* Agricultural Research Council Unit of Animal Genetics, Institute of Animal Genetics, Edinburgh.

say, if a mouse sheds fewer than average on one side it is likely to shed more than average on the other. This fact has been noted by several investigators who have counted mouse eggs or corpora lutea (e.g. Danforth & de Aberle, 1928; Hollander & Strong, 1950; Bowman & Roberts, 1958), and it is also true of rabbits (Adams, 1959) and of guinea-pigs (Eckstein & McKeown, 1955). In contrast, there is a positive correlation between the numbers of eggs shed by the two ovaries of mice after induced superovulation (Fowler & Edwards, unpublished). No one seems to have examined the nature of the correlations, nor considered what light they may throw on the physiological mechanisms of ovulation. This is the purpose of the present paper.

The numbers of eggs shed at one oestrus have been shown to be randomly distributed between the two ovaries in mice and several other species (see Brambell, 1956). We shall examine the distribution between the two ovaries of mice in more detail than has been done before, and consider how the correlation between sides is related to this distribution after both natural ovulation and induced superovulation. If the distribution of eggs is random, the numbers shed by each ovary will conform to the expectations of a binomial distribution. A binomial distribution with respect to the numbers of eggs shed by the left and right ovaries respectively, means that each egg has a certain probability of having been shed by, say, the left ovary, and this probability is the same for all the eggs shed at that oestrus; whatever the physiological mechanism is that determines the total number of eggs shed at any one oestrus, the ovary from which each egg is shed is a matter purely of chance.

SOURCES OF DATA

Natural mating was judged by the presence of a vaginal plug. Eggs were counted on the morning that the vaginal plug was found, by dissection of the Fallopian tubes. Eggs from mice given superovulation treatments were counted in the same way, ovulation having been induced by an intraperitoneal injection (on the right side) of varying doses of pregnant mares' serum (PMS) followed after 40 hr by human chorionic gonadotrophin (HCG) (Fowler & Edwards, 1957). Corpora lutea were counted by dissection of the ovaries of pregnant females, most of them at 16 to 17 days after the vaginal plug, but a few at earlier stages. All mice were adult and were between 6 and 14 weeks of age.

The data comprise 697 mice counted for eggs and 390 counted for corpora lutea. The mice came from nine distinct outbred strains and some partially inbred and crossbred mice derived from them. The designations and characteristics of the strains are as follows:

NF, NS, NC: selected for large body weight, for small body weight and an unselected control, respectively, all derived from the same base (Falconer, 1953, 1955).

CFL, CFS, CRL: selected respectively for high and low 3- to 6-week growth on normal diet and for high growth on restricted diet, all derived from the same base (Falconer, 1960a).

JH, JL, JC: selected for large and small litter size and an unselected control,

all derived from the same base as the three 'C-strains' above (Falconer, 1955, 1960b).

In addition there were the following miscellaneous groups, all derived from the 'J-strains' or the same base, designated for reference in this paper as follows:

JB: crosses between partially inbred lines.

JR: various partially inbred and crossbred mice. 'Series I' of Falconer & Roberts (1960).

JF: various partially inbred and non-inbred mice. 'Series II and III' of Falconer & Roberts (1960).

Where the strains are referred to separately in the Tables, the generations from which the mice were derived are given in brackets, and the observer who did the counting is indicated by initials. The doses of PMS given to induce superovulation are also shown.

METHODS OF ANALYSIS AND RESULTS

COMPARISON OF LEFT AND RIGHT OVARIES

It is necessary first to find out if the two ovaries shed on the average the same number of eggs, or if there is any overall bias towards the left or right side. Only some of the data can be used for this purpose because in the remainder, though the two ovaries were recorded separately, the left and right sides were not distinguished. The data given in Table 1 come from natural mating in JH, JL

TABLE 1

COMPARISON OF LEFT AND RIGHT OVARIES

	No. mice	*No. eggs or corpora lutea*		
		Left	*Right*	*% Left*
Natural ovulation				
Eggs	159	801	841	48·78
Corpora lutea	286	1434	1516	48·61
Superovulation				
Eggs	49	306	325	48·49
Total	494	2541	2682	48·65

and JC (Generations 16, 17, 32, 33), JR and JF. Differences between strains were trivial and the strains are combined in the table. The counts of eggs and of corpora lutea are shown separately, though they gave almost exactly the same result. There is also a smaller amount of data from egg counts in the NF, NC and CFS strains following superovulation, and these gave the same result as natural ovulation. Taken all together, the data show a very slight bias in favour of the right ovary, which yielded 51·65% of all eggs or corpora lutea. The χ^2 testing deviation from 50% is 3·81, which is on the border of significance at the 5% level. The slight bias is therefore probably real. But even if real, the difference from equality is too small to make any appreciable difference to the expectations based on equality. In the analyses that follow, therefore, the overall ratio of left to right is taken to be 50%.

BINOMIAL DISTRIBUTION OF EGGS BETWEEN THE TWO OVARIES

Expected and observed distributions in the most frequent classes

Since there is no material bias towards left or right we can now disregard the distinction between left and right, and deal with the distributions in terms of the difference between one side and the other. For example, two mice, one with three on the left and five on the right, the other with five on the left and three on the right, are equivalent; both have a difference of two. We may first examine the distributions of these differences among the commonest total egg numbers and see how they agree with the binomial expectation. The commonest total numbers were eight to eleven, each of which is represented by more than fifty mice in the total data on egg counts. The binomial expectation for the differences can be found from statistical tables (e.g. the *Tables* cited in the reference list). For example, among mice having a total of eight eggs, 27% are expected to have no difference (i.e. four on both sides), 44% are expected to have a difference of two (i.e. 3 : 5 or 5 : 3), 22% to have a difference of four (2 : 6 or 6 : 2), 6% a difference of six (1 : 7 or 7 : 1), and 1% a difference of eight (nil on one side and eight on the other). Table 2 shows the expected and observed

TABLE 2

OBSERVED DISTRIBUTIONS OF DIFFERENCES BETWEEN SIDES FOR COMPARISON WITH THE BINOMIAL EXPECTATIONS

Total No. eggs	*No. mice*		*Difference between sides*						*Mean squared difference*
			0	*2*	*4*	*6*	*8*	*10*	
8	55	Exp.	15·0	24·1	12·0	3·4	0·4	–	8
		Obs.	12	22	15	5	1	–	10·4
10	83	Exp.	20·4	34·0	19·5	7·3	1·6	0·2	10
		Obs.	19	26	27	8	3	0	12·2
			1	*3*	*5*	*7*	*9*	*11*	
9	93	Exp.	45·8	30·5	13·1	3·3	0·4	–	9
		Obs.	41	31	18	2	1	–	10·2
11	58	Exp.	26·2	18·7	9·3	3·1	0·6	0·1	11
		Obs.	29	17	11	1	0	0	8·7

The figures in the body of the table are numbers of mice expected and observed, the observed numbers coming from all the data on egg-counts combined.

numbers of mice according to the difference between sides. For example, there were fifty-five mice with totals of eight eggs, of which twelve mice had no difference between sides; the expected number with no difference is 27% of fifty-five, which is fifteen. The agreement between the observed and expected numbers is obviously very close.

In order to obtain a comprehensive test of the agreement with the binomial distribution, it is necessary to condense the distribution of differences between sides into a single figure for comparison with expectation. This can be done by making use of the property of a binomial distribution, that the mean of the squared differences is expected to equal the total number. This property can be deduced in the following manner. It is well known that the variance of the numbers observed in one of two binomially distributed classes is npq, a formula that can be found in any statistical text. Applied to the present problem, n is the total number of eggs from both ovaries together; p and q are both $\frac{1}{2}$, as shown in the previous section, so the expected variance reduces to $\frac{1}{4}n$. The 'variance' is, by definition, the average squared deviation from the mean number. The deviation of each side is half the difference between the two sides. So we can also express the expected variance in terms of the difference between sides, d, as the average of $(\frac{1}{2}d)^2$. Therefore, if N is the number of mice examined, all with the same total number of eggs, n, the expectation is that $\Sigma\,(\frac{1}{2}d)^2/N = \frac{1}{4}n$, or $\Sigma\,d^2/N = n$. That is, the mean of the squared differences is expected to equal the total, if the distribution is binomial.

The mean of the squared differences found among the mice with total egg numbers of eight to eleven are shown at the right of Table 2. Three of them are a little in excess of expectation and one is below. Whether the discrepancies are significant or not will be considered later.

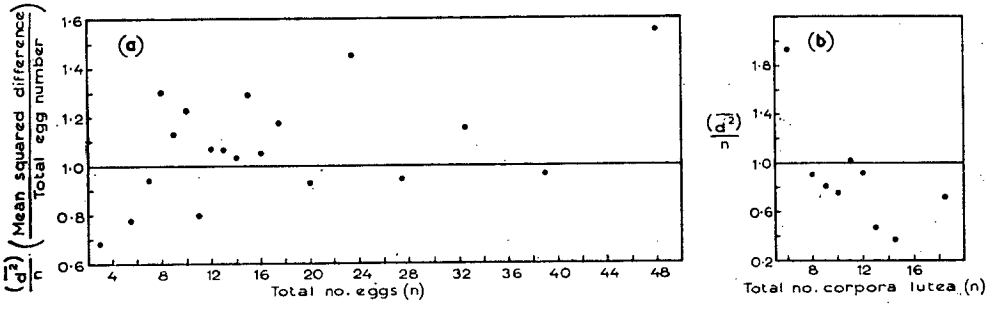
Expected and observed distributions in all classes

The first point to be examined with the mean squared difference is whether the inequality of the numbers of eggs shed by the two ovaries varies according to the total number of eggs. It might be expected that there would be a limit to the number of eggs that one ovary could shed at one time. This would tend to reduce the average difference between sides when large numbers of eggs were ovulated, particularly with superovulation. For the purpose of this analysis all the data on egg counts from all stocks and treatments were put together. Similarly, all the data on corpora lutea counts were put together, but not combined with the egg data. The mean squared difference between sides was then computed for all mice having the same total number of eggs. These total egg numbers ranged from two to fifty-four, but the lower and higher numbers were represented by too few mice to give reliable comparisons between the observed and expected mean squared difference. The data were therefore grouped as follows. For each total egg number the mean squared difference between sides was divided by the total egg number: this ratio $(\overline{d^2}/n)$ will be unity if the distribution is binomial. The ratios of the classes to be grouped together were then averaged, with weighting according to the numbers of mice. The ratios obtained are shown in Text-fig. 1, from which the following conclusions can be drawn:

(1) The inequality between sides is not less when the number of eggs is large, and the distribution between sides is approximately binomial over the whole of the range of egg numbers. There is, perhaps, a suggestion that the inequality tends to increase with the higher egg numbers, but this tendency, when tested by a regression analysis, was not significant.

(2) With corpora lutea counts the inequality between sides tends to decrease as the total number increases. The first point, representing mice with five to seven corpora lutea, may be spuriously high: it included four mice with all the corpora lutea in one ovary and none in the other — one with five, one with six and two with seven. Even if this point is included, however, the tendency to decrease is not significant. (The computed regression is: 0.057 ± 0.036.)

(3) Most of the points representing egg counts show a difference between sides greater than expectation, and most of the points representing corpora lutea counts show a difference less than expectation. The next point to



TEXT-FIG. 1. *Ratio of mean squared difference between sides, to total egg number (left + right), plotted against the total egg number.* The binomial expectation for this ratio is unity. (a) Egg-counts, all data combined. (b) Corpora lutea counts.

be examined must, therefore, be the overall agreement with the binomial expectation.

Significance tests on departures from expected ratios

In order to test the overall agreement with the binomial expectation, it is necessary to condense the measure of the inequality between sides still further by combining the mean squared difference from all mice irrespective of the total number of eggs. This will also enable us to compare strains and treatments. The measure of inequality that seems most appropriate is again the ratio of the mean of squared differences to the total number of eggs, which has an expectation of 1. Combination of different totals has been done by summation of all squared differences and division of this by the sum of all totals, i.e. $\Sigma d^2/\Sigma n$. This ratio weights the measure by the number of eggs counted rather than by the number of mice examined. The ratios obtained for the different strains and treatments are given in Tables 3 and 4.

The test of significance of the departures from the expected value of 1 for this ratio ($\Sigma d^2/\Sigma n$) was made by a method described by Robertson (1951). The basis of this method is an estimate of the heterogeneity between mice in their left-right difference, which may be explained in words as follows. The mean proportion of the eggs shed at one ovulation that come from one of the ovaries, say the left, is $\frac{1}{2}$ (very nearly). Individual mice, however, vary widely round this mean ratio. How much of this variation represents real differences between the mice, over and above the chance variation expected from a binomial distribution? This real variation between mice is the 'heterogeneity' that the method

TABLE 3

ANALYSES OF EGG COUNTS AND CORPORA LUTEA COUNTS FOLLOWING NATURAL OVULATION, CLASSIFIED ACCORDING TO THE STRAIN

Strain	*Observer*	*No. mice*	*Mean No. eggs* ($\bar{n}$)	*Variance of total egg No.* (σ^2_n)	$\frac{\Sigma d^2}{\Sigma n}$	*Heterogeneity* ($H \times 10^4$)	*Standard error of H* ($\sigma_H \times 10^4$)	H/σ_H	*Correlation betw. sides*	
									Obs.	*Exp.*
Egg counts										
NF (36 to 40)	R.E.F.	33	11·1	2·86	1·06	14·4	57·3	0·25	−·61	−·59
NS (29 to 32)	R.E.F.	13	4·9	1·23	0·54	−291·7	228·2	−1·28	−·35	−·60
NC (21 to 25)	R.E.F.	24	7·5	2·87	0·89	−40·5	100·6	−0·40	−·40	−·45
CFL (17 to 19)	R.G.E.	21	14·9	2·13	0·65	−63·0	53·5	−1·18	−·64	−·75
CFS (17 to 20)	R.G.E.	29	8·9	2·50	1·28	84·8	76·7	1·10	−·64	−·56
CRL (17 to 20)	R.G.E.	23	16·0	5·09	0·97	− 5·3	47·1	−0·11	−·51	−·52
JH (18 to 19)	R.G.E.	27	11·7	4·62	1·57	129·3	59·9	2·16*	−·60	−·43
JH (16 to 17)	D.S.F.	45	10·4	2·20	0·85	−37·8	52·7	−0·72	−·60	−·65
JH (32)	D.S.F.	30	13·7	4·22	1·73	139·5	48·4	2·88*	−·70	−·53
JL (18 to 19)	R.G.E.	25	10·1	2·03	1·09	23·5	73·1	0·32	−·69	−·67
JL (16 to 17)	D.S.F.	34	8·5	1·39	1·22	72·3	75·2	0·96	−·77	−·73
JL (32)	D.S.F.	20	10·3	2·22	1·30	79·2	79·9	0·99	−·72	−·65
JC (32)	D.S.F.	30	8·9	2·13	1·66	210·1	75·6	2·78*	−·75	−·61
All		354	10·61	2·736	1·175	41·7	17·8	2·34*	−·64	−·59
Corpora lutea										
JH (33)	D.S.F.	38	13·1	5·79	0·53	−94·1	44·1	−2·13*	−·09	−·39
JL (33)	D.S.F.	67	9·9	3·09	0·95	−12·2	45·1	−0·27	−·51	−·53
JC (33)	D.S.F.	37	9·1	1·52	0·95	−16·3	67·2	−0·24	−·70	−·71
JR	R.C.R.	144	10·1	3·76	0·93	−18·7	30·2	−0·62	−·43	−·46
JF	D.S.F.	104	11·0	7·19	0·65	−82·8	32·2	−2·57*	+·01	−·21
All		390	10·48	4·550	0·81	−48·0	17·5	−2·75*	−·30	−·40

* Heterogeneity significantly different from zero: $P = 0{\cdot}05$ or less.

The figures in brackets after the strain designations refer to the generation from which the mice came. The combined values entered for 'all' strains are weighted averages.

estimates. Alternatively, the variation of the ratios actually found might be less than would be expected on a chance basis. In that case, the 'heterogeneity' would be negative. The method provides a standard error for the estimate of heterogeneity, from which the significance of differences from the binomial expectation can be assessed. The heterogeneity, H, is estimated as follows:

$$H = \frac{2\,\Sigma(d^2-n)}{8\,\Sigma n(n-1)},$$

where d is the difference between sides and n is the total number of eggs from each mouse. We have already seen that, if the distribution is binomial, d^2 will equal n. Therefore an excess of inequality between sides will appear as a positive heterogeneity; and if, conversely, the number of eggs from the two ovaries are more alike than would be expected by chance, this will appear as a negative heterogeneity. The sampling variance of the estimate of the heterogeneity is

$$\sigma^2_H = \frac{1}{8\,\Sigma n(n-1)}$$

and the standard error is the square root of this.

The estimates of heterogeneity and their standard errors found in the different strains and treatments are listed in Tables 3 and 4. The three main groups of

TABLE 4

ANALYSES OF EGG COUNTS FOLLOWING SUPEROVULATION, CLASSIFIED ACCORDING TO THE STRAIN OR THE DOSAGE OF PMS

Strain or dosage	*No. mice*	$\frac{\Sigma d^2}{\Sigma n}$	*Heterogeneity* ($H \times 10^4$)	*Standard error of H* ($\sigma_H \times 10^4$)	H/σ_H	*Correlation betw. sides*	
						Obs.	*Exp.*
NF (36-40)	54	0·99	— 1·7	29·3	−0·06		
NS (29-32)	45	1·58	137·5	52·4	2·62*		
NC (22-26)	45	0·75	— 28·6	28·8	−0·99		
CFL (17-20)	49	1·34	31·8	21·5	1·48		
CFS (16-20)	24	0·62	— 48·9	44·8	−1·09		
JH (18-19)	25	0·96	— 4·3	30·0	−0·14		
JL (18-19)	19	1·22	19·4	31·5	0·62		
JB	82	1·06	6·7	18·9	0·36		
All	343	1·08	8·9	9·8	0·90		
¼ i.u.	9	1·22	77·8	166·7	0·47	+·03	+·13
½ i.u.	62	1·20	47·1	43·6	1·08	−·05	+·04
1 i.u.	109	0·99	— 1·1	26·0	−0·04	+·45	+·45
3 i.u.	147	1·02	2·2	12·2	0·18	+·63	+·63
6 i.u.	16	1·49	33·5	25·7	1·30	+·59	+·71

* Heterogeneity significantly different from zero: $P = 0{\cdot}05$ or less.
All observations by R.G.E. or R.E.F.

mice in the data show different results. Egg counts following natural ovulation show a significant excess of inequality between sides over and above what would be expected from a binomial distribution. The overall ratio of the mean squared

difference between sides to the total number of eggs ($\Sigma d^2/\Sigma n$) is 1·175 instead of the expected 1·0. The heterogeneity variance between mice is 2·34 times its standard error, and so the excess of inequality, though very small, is significant at the 2% level. Corpora lutea counts following natural ovulation show, in contrast, significantly less variation between sides than would be expected by chance. The ratio $\Sigma d^2/\Sigma n$ is 0·808 and the (negative) heterogeneity variance is 2·75 times its standard error, which is significant at the 1% level. Since egg counts and corpora lutea differ significantly from expectation in opposite directions it is clear that they differ from each other significantly. The difference of heterogeneity between the two sets of data is 3·6 times its standard error and is significant at the 0·1% level. Finally, the egg counts after superovulation show a slight but non-significant excess of variation between sides. The ratio $\Sigma d^2/\Sigma n$ is 1·08, and the heterogeneity variance is less than its standard error. The difference of heterogeneity between the naturally ovulating and superovulated mice is not significant: it is 1·6 times its standard error and this has a

TABLE 5

COMPARISONS OF STRAINS AND OBSERVERS IN RESPECT OF THE ESTIMATES OF HETEROGENEITY

	d.f.	χ^2	*P*
Natural ovulation			
Egg counts: between strains	12	22·2	·05-·02
between observers*	1	3·6	·1-·05
Corpora lutea: between strains	4	4·1	·5-·3
Superovulation			
Between strains	7	10·9	·2-·1
Between doses	4	2·3	·7-·5

* (D.S.F. *vs* R.E.F. & R.G.E.)

probability of 10%. We cannot conclude, therefore, that the inequality between sides is really less in superovulated than in naturally ovulating mice.

The range of estimates of heterogeneity obtained from the separate strains may seem from an examination of Table 3 to be rather large; the estimates range from −1·28 to +2·88 times their individual standard errors. Also, there seems possibly to be a difference between observers, R. E. F. and R. G. E. obtaining low heterogeneities and D. S. F. high ones. Are these differences between strains and between observers, in respect of the estimates of hereogeneity, statistically significant? Table 5 shows χ^2 tests of these differences, the χ^2 values being obtained as follows. If the ratios of H/σ_H (given in Table 3) are normally distributed, then the squares of the ratios will be distributed as χ^2. Summation of $(H/\sigma_H)^2$ over, say, n strains gives a total χ^2 with n degrees of freedom. Subtraction of the value of $(H/\sigma_H)^2$ for all strains combined gives a χ^2 with $(n\text{-}1)$ degrees of freedom which tests the significance of differences in H between strains. Table 5 shows that the differences between strains in egg counts following natural ovulation are significant at the 5% level, and the differences between observers approach significance at this level. Observers

and strains are, however, confounded in the analysis, and the apparent difference between observers could well arise from the differences between strains, since only two strains were common to both observers.

The conclusions from the foregoing analyses are that the difference between the numbers of eggs shed from the two ovaries is on the average a little greater than would be expected by chance, but the difference between the corpora lutea counted in the two ovaries is less than would be expected by chance and less than that of the eggs shed. The deviations from chance expectation are, however, very small and the distribution of eggs between the two ovaries is very close to the binomial expectation.

CORRELATION BETWEEN LEFT AND RIGHT SIDES

The starting point for the present study was the correlation between the numbers of eggs from the two ovaries. How are the observed correlations related to the binomial distribution which has been demonstrated? The sign and magnitude of the correlation must depend on the relative magnitude of the variation between sides and the variation of total egg number between individual mice. Thus, if all mice shed the same total number of eggs there would be perfect negative correlation between sides; if, on the other hand, there were no differences between the two sides but mice varied in total egg number, then there would be perfect positive correlation between sides. The object of this section, therefore, is to deduce the theoretically expected correlation in terms of these two sources of variation and then to show that the observed correlations follow the expected pattern.

The correlation coefficient, r, can be expressed in terms of the variance of the total egg number, n and the variance of the difference between sides, d, thus:

$$r = \frac{\sigma_n^2 - \sigma_d^2}{\sigma_n^2 + \sigma_d^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

This expression can be derived as follows. Let L and R be the number of eggs on the left and right sides of any mouse. Let $d = L-R$, the sign of the difference now being taken into account. Then $L = \frac{1}{2}(n+d)$ and $R = \frac{1}{2}(n-d)$. The sums are $\Sigma L = \frac{1}{2}(\Sigma n+\Sigma d)$, $\Sigma R = \frac{1}{2}(\Sigma n-\Sigma d)$, and the sum of products is $\Sigma LR = \frac{1}{4}\Sigma(n+d)(n-d) = \frac{1}{4}(\Sigma n^2-\Sigma d^2)$. The covariance is

$$\mathrm{cov}_{LR} = \frac{1}{N-1}\left(\Sigma LR - \frac{\Sigma L\,\Sigma R}{N}\right),$$

where N is the number of mice. Writing L and R in terms of n and d gives

$$\begin{aligned}\mathrm{cov}_{LR} &= \frac{1}{N-1}\left[\tfrac{1}{4}(\Sigma n^2-\Sigma d^2) - \tfrac{1}{4}\frac{(\Sigma n)^2-(\Sigma d)^2}{N}\right] \\ &= \tfrac{1}{4}\cdot\frac{1}{N-1}\left\{\left[\Sigma n^2 - \frac{(\Sigma n)^2}{N}\right] - \left[\Sigma d^2 - \frac{(\Sigma d)^2}{N}\right]\right\} \\ &= \tfrac{1}{4}\,(\sigma_n^2 - \sigma_d^2).\end{aligned}$$

The variance of the numbers on the two sides will be equal, so $\sigma_L\sigma_R = \sigma_L^2 = \sigma_R^2$ = variance of $\frac{1}{2}(n \pm d) = \frac{1}{4}(\sigma_n^2+\sigma_d^2 \pm 2\ \mathrm{cov}_{nd})$. The covariance terms will

cancel out because half will be positive and half negative and of equal magnitude. Thus $\sigma_L\sigma_R = \frac{1}{4}(\sigma_n^2+\sigma_d^2)$. The correlation between left and right sides is

$$r_{LR} = \frac{\mathrm{cov}_{LR}}{\sigma_L\sigma_R} = \frac{\sigma_n^2-\sigma_d^2}{\sigma_n^2+\sigma_d^2}.$$

The above expression for the correlation rests on no assumption other than that the variances of the numbers shed by left and by right ovaries are equal, which is a reasonable supposition to make. This formula was used to calculate the observed correlations, and these are given in Tables 3 and 4. Table 3 refers to natural ovulation and the correlations found in each strain are given. Those based on egg counts are all negative and range from −0·35 to −0·77; those based on corpora lutea counts range from +0·01 to −0·70. Table 4 refers to superovulation and the correlations found after each dosage level are given. They are about zero after the two low dosages and are positive after the higher dosages, the highest value being +0·63.

If we apply the conclusion, arrived at in the previous section, that the distribution of eggs between the two sides is binomial, or nearly so, we can arrive at an expression for the correlation between sides in terms of the variance and mean of the total egg number. Since the mean number of eggs shed by left and by right ovaries is equal ($\bar{d} = 0$) it follows that the variance of the difference, σ_d^2, is equal to the mean squared difference, $\Sigma d^2/N$. This, as was explained earlier, is equal to the total number of eggs, n, if the distribution between sides is binomial. We can therefore substitute the mean of the total egg number, $\bar{n}$, for the variance of the difference, σ_d^2, in equation (1), and write the correlation in the form

$$r = \frac{\sigma_n^2-\bar{n}}{\sigma_n^2+\bar{n}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

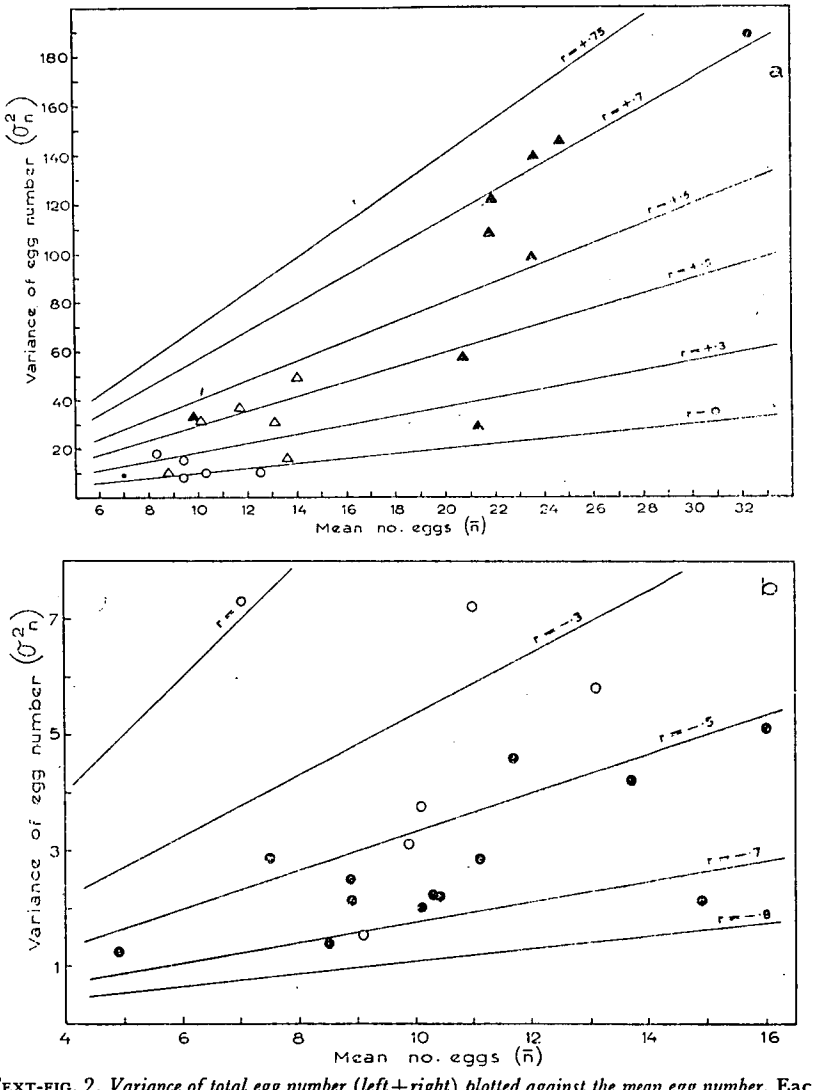
This formula shows that, with a binomial distribution between sides, the correlation will be negative when the variance is less than the mean, and positive when the variance exceeds the mean.

The expected correlations, calculated from Equation 2, are given in Tables 3 and 4. The observed correlations deviate from the expected correlations only in so far as the mean squared difference between sides deviates from the total egg number. The close correspondence between the observed and expected correlations therefore provides no new information: it merely reflects the close agreement with the binomial distribution which has already been established.

The way in which the expected correlation depends on the variance and the mean may be more easily appreciated from the graphical representation in Text-fig. 2. This shows each strain and treatment plotted according to its variance and mean of total egg number. The superovulation data are here subdivided into both strains and treatments, as in Table 6, so that there is a point for each strain at each dosage level with which it was treated. Some of these points are based on rather few animals and they are consequently rather widely scattered. The correlations expected are related to the positions of points on the graphs by Equation 2, and straight lines are drawn to mark the positions of various levels of the correlation coefficient. The positions of the

points in relation to these lines show how the expected correlations vary from natural ovulation to superovulation, and with increasing dosages of PMS.

The fact that the correlations change from negative in naturally ovulating mice to positive after superovulation is purely the consequence of the relatively



TEXT-FIG. 2. *Variance of total egg number (left+right) plotted against the mean egg number.* Each point refers to a strain or treatment. The straight lines represent various values of the expected correlation between left and right ovaries, as explained in the text. (a) Superovulation. Dosage: • = ¼ i.u.; ○ = ½ i.u.; △ = 1 i.u.; ▲ = 3 i.u.; ● = 6 i.u. (b) Natural ovulation (● = egg counts; ○ = corpora lutea counts).

greater variance of total egg number after superovulation, together with the binomial distribution between sides. This effect of superovulation on the variance may be of physiological interest, and it will now be examined in more detail.

TABLE 6

MEANS AND VARIANCES OF TOTAL EGG NUMBERS FOLLOWING SUPEROVULATION, CLASSIFIED BY BOTH STRAINS AND DOSAGE

	Dosage of PMS (i.u.)														
	¼			½			1			3			6		
Strain	N	$\bar{n}$	σ^2_n	N	$\bar{n}$	σ^2_n	N	$\bar{n}$	σ^2_n	N	$\bar{n}$	σ^2_n	N	$\bar{n}$	σ^2_n
NF	–	–	–	17	9·4	8·2	17	14·0	49·0	20	20·7	56·9	–	–	–
NS	9	7·0	9·0	14	10·3	10·1	17	10·1	31·0	5*	9·8	33·2	–	–	–
NC	–	–	–	9	9·4	14·8	21	11·7	36·9	15	24·7	145·1	–	–	–
CFL	–	–	–	15	12·5	10·3	9	13·6	15·5	9	21·3	28·8	16	32·3	187·8
CFS	–	–	–	7	8·3	18·2	8	8·8	10·2	9	21·9	121·6	–	–	–
JH	–	–	–	–	–	–	–	–	–	25	21·8	107·8	–	–	–
JL	–	–	–	–	–	–	–	–	–	19	23·6	138·8	–	–	–
JB	–	–	–	–	–	–	37	13·1	30·4	45	23·5	97·8	–	–	–
All	9	7·0	9·0	62	10·21	11·12	109	12·21	32·11	147	22·26	99·53	16	32·3	187·8

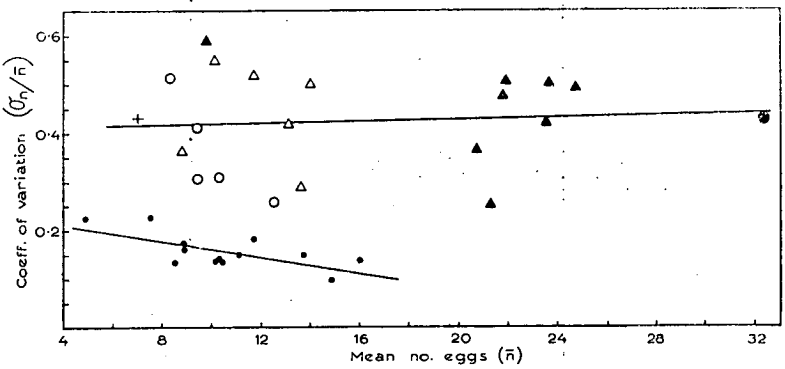
* Excluding seven mice that gave no eggs.

All observations by R.G.E. or R.E.F. N = number of mice; $\bar{n}$ = mean number of eggs (left+right); σ^2_n = variance of egg number. The combined values at the foot are weighted averages.

DOSAGE EFFECT IN SUPEROVULATION

Consideration of the variance of total egg number and its relation to the mean egg number leads to a conclusion that is of interest in connection with superovulation. Text-fig. 2 shows that the variance increases as the mean increases, both with superovulated and naturally ovulating mice. (In this section, 'variance', 'mean' and 'egg number' all refer to total egg number, i.e. the sum of the numbers from the two ovaries). Text-fig. 2 shows also that the variance is very much greater after superovulation than after natural ovulation (see above). Is this increased variance consistent with the greater mean, or does the superovulation itself cause a greater variability?

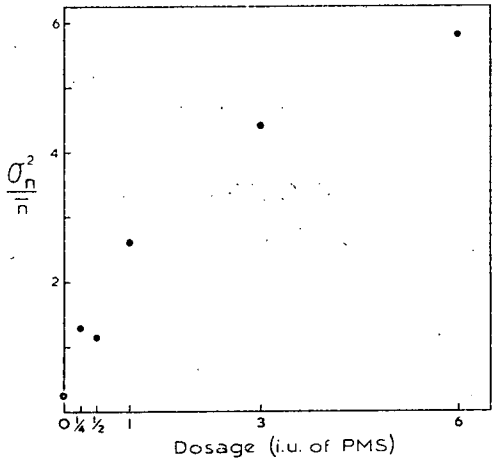
The answer to this question can be seen most clearly from a consideration of the coefficient of variation, i.e. the ratio of standard deviation to mean egg number ($\sigma_n/\bar{n}$), though as will be seen later this is not the most appropriate measure of variation for making the comparison. The coefficients of variation are plotted against the mean egg numbers in Text-fig. 3 from the data contained



TEXT-FIG. 3. *Relationship between coefficient of variation and mean egg number in different strains and after different dosages of* PMS. • = natural; + = ¼ i.u.; ○ = ½ i.u.; △ = 1 i.u.; ▲ = 3 i.u.; ● = 6 i.u.

in Tables 3 and 6. Each point represents a strain, and the different dosages used to induce superovulation are shown separately. Over the range of mean egg numbers covered by natural ovulation, where a direct comparison can be made, the coefficient of variation is considerably higher after superovulation than after natural ovulation. Superovulation therefore causes an additional source of variation not present in natural ovulation, even when the mean egg numbers are the same. Another conclusion to be drawn from Text-fig. 3 is that the coefficient of variation of egg number after natural ovulation is not constant, but declines as the mean increases. This trend is statistically significant at the 1% level. The regression of the coefficient of variation on the mean egg number is $-0{\cdot}0082 \pm 0{\cdot}0025$ ($t_{[11]} = 3{\cdot}3$; $P = 0{\cdot}01$). After superovulation, in contrast, the coefficient of variation does not decline as the egg number increases. The regression is $+0{\cdot}0005 \pm 0{\cdot}003$, and the two regression coefficients are significantly different from each other at the 5% level. Thus, not only is the variation relatively greater after superovulation than after natural ovulation, but the form of the relationship between the variation and the mean is different. The

interpretation of this different relationship is complicated by the fact that the differences in mean egg number following superovulation are associated with differences of dosage, and the amount of variation may be influenced by the dosage level, as distinct from the mean number of eggs shed. If the relationship is examined within each of the three intermediate dosage levels, which are represented by several strains, it is found that there is a tendency for the coefficient of variation to decline with increasing mean, just as with natural ovulation. The mean regression, within dosage levels, of the coefficient of variation on the mean egg number is $-0{\cdot}0107 \pm 0{\cdot}0073$, which is nearly the same as with natural ovulation. The regression is, however, not significantly different from zero, and it can only be taken as a suggestion that the coefficient of variation may follow the same relationship with the mean within any one dosage level as it does in naturally ovulating mice. The empirical fact demonstrated by the graph is that the coefficient of variation is the same at all dosage levels, whereas it is less with high natural egg numbers than with low.



TEXT-FIG. 4. Ratio of variance to mean egg number in relation to dosage of PMS.

Since the coefficient of variation is not constant in naturally ovulating mice, this measure of variation is not a suitable one for the comparison of variation in groups with different mean egg numbers. A suitable measure is the ratio of variance to mean ($\sigma^2_n/\bar{n}$). This ratio was found to remain constant over the range of mean egg numbers for natural ovulation in the data. The regression of the ratio on the mean egg number was $+0{\cdot}0003 \pm 0{\cdot}0077$. The ratio of variance to mean therefore seems to provide the more appropriate measure of variation for an assessment of the effects of dosage level on the variability of egg number following superovulation. The ratio of variance to mean is plotted in Text-fig. 4 against the dosage. The ratio for each dosage level is the overall mean of the ratios for each strain tested at that dosage, weighted by the number of mice. The ratio for natural ovulation is shown against a dosage of 0. If the variation after superovulation behaved as it does after natural ovulation, and if the only effect of higher doses was to produce a higher mean egg number, then the

ratio of variance to mean would be expected to remain constant over all doses and to have a value equal to that after natural ovulation. Text-fig. 4 shows clearly that it is above the level for natural ovulation at the lowest doses and that it increases very much at the higher doses. The dosage of PMS required to induce a mean ovulation rate equal to the natural rate was calculated for the N- and C-strains used in this work by Fowler & Edwards (1960) from regressions of total egg number on dosage. The values obtained were below $\frac{1}{4}$ i.u. for the NS (small) strain and between 0·4 and 1·2 i.u. for the other N- and C-strains. At dosages of $\frac{1}{4}$ to 1 i.u., the variance is five to ten times as great as after natural ovulation. At a dose of 3 i.u., the variance is twenty-four times as great as it would be if the same mean egg number were produced by natural ovulation. Thus the induction of superovulation by PMS introduces an additional source of variation which is not present in naturally ovulating mice, and this additional variation increases with higher doses of PMS.

The conclusions about the relationship between variation and mean egg number which have been drawn in this section are that the variance is proportional to the mean egg number after natural ovulation, and possibly also after superovulation when the dose is not varied; but when superovulation is induced by varying doses of PMS then the standard deviation is proportional to the mean egg number corresponding to each dose.

DISCUSSION

The essential requirement, in statistical terms, for any postulated physiological mechanism of ovulation to be consistent with the random (binomial) distribution of eggs between the two ovaries, is that the maturation of one follicle reduces the probability of the maturation of any other follicle, whether in the same ovary or in the other. This requirement can only be met if the maturation of any follicle changes the conditions influencing both ovaries. One obvious way of meeting this requirement is that the maturation of each follicle uses up some of the circulating hormone; less would then be available for the stimulation of other follicles to maturation, whether in the same ovary or in the other. Whatever the physiological basis it may well be common to most mammals. A random distribution of corpora lutea between the two ovaries exists in the common shrew (Brambell, 1935), the lesser shrew (Brambell & Hall, 1937) and the bank vole (Brambell & Rowlands, 1936). The negative correlation between ovaries reported in guinea-pigs (Eckstein & McKeown, 1955) and rabbits (Adams, 1959) also makes it probable that the same binomial distribution may occur in most mammals including women; in species ovulating one egg, a random ovulation would often give the impression of regular alternation (Brambell, 1956). Nevertheless, in several species of mammals the female has been reported to ovulate more frequently from one ovary, e.g. the mare, the wild mountain viscacha and some species of bats (Eckstein & Zuckerman, 1956).

The random distribution between sides, even in mice with the largest numbers of eggs, proves that if there is a limit to the number of follicles that can mature in one ovary at the same oestrus, this limit has not been reached even

with superovulation by the highest doses of PMS administered. This conclusion is supported by three other lines of evidence. Edwards & Fowler (1960) found that a second ovulation almost always resulted from a second treatment with PMS and HCG given between 1 and 3 days after the first treatment. Jones & Krohn (1961) showed, from detailed counts of the numbers of oocytes in the ovaries of inbred and hybrid mice of various ages between birth and senescence, that many hundreds of oocytes were available in all strains at ages comparable to those of the mice used in the present experiment. Lastly, after complete removal of one ovary and even the partial removal of the other, the remaining ovarian tissue hypertrophies and can shed approximately the same number of eggs at each oestrus as the two ovaries had done previously (e.g. Hollander & Strong, 1950). But the number of oocytes available in these ovarian fragments declines with an increasing interval after removal of the ovarian tissue (e.g. Lipschütz, 1928; Mandl, Zuckerman & Patterson, 1952). The largest number of eggs counted from one ovary in the present data was thirty-one, which occurred in three different mice.

Statistical analysis showed that the variation between the numbers of eggs counted from the two ovaries after natural ovulation was a little greater than the random amount, but the variation in corpora lutea between the two ovaries was less than the random amount. The following possible causes of increased or of reduced variation between sides may be suggested:

(i) Some mice might be slightly 'left-sided' and others slightly 'right-sided'. This would increase the variation both of egg counts and of corpora lutea counts. A difference of this sort between the two ovaries might result from a difference in the hormonal supply, perhaps through a difference in the blood supply to the two ovaries, or from the impairment of one ovary by disease in some mice.

(ii) The occurrence of follicles not represented by an egg in the Fallopian tube. Two processes leading to such an event are the formation of corpora lutea atretica, i.e. the luteinization of a follicle without the liberation of its oocyte, and intraovarian ovulation, in which the oocyte is liberated into the ovarian tissue. A first report of the latter process and measurements of the incidence of both processes in mice have been given by Jones & Krohn (1961). The loss of eggs, from any cause, would increase the variation between sides in egg counts, but not in corpora lutea counts. The occurrence of atretic corpora lutea would do the same, unless they require less hormone during their formation than do normal corpora lutea. Both processes would, however, reduce the variation between sides if they occurred relatively more frequently in ovaries that shed a large number of eggs. They would then tend to diminish high counts without altering the low counts.

(iii) The occurrence of polyovular follicles; i.e. the presence of more than one oocyte in a follicle. The incidence of polyovular follicles in 6-week-old 'Swiss' mice has recently been estimated as about 10% (Kent, 1960), though it declined in older mice. An incidence as high as this in our material could have had some effect on the variation between the two ovaries, though it is not clear to what extent polyovular follicles contribute to the eggs ovulated. Allen, Brambell & Mills (1947) showed that the incidence of eggs from these follicles

was unimportant (0·23%) in the wild rabbit. If a polyovular follicle requires the same amount of hormone for its maturation as a normal follicle, then the variation between sides in egg counts would be increased but that of corpora lutea counts would not. If the amount of hormone required is in proportion to the number of eggs maturing, then polyovular follicles would have no effect on the variation between sides in egg counts, but would reduce the variation in corpora lutea counts. Thus the occurrence of polyovular follicles could account for either the excess variation between sides in egg counts or the reduced variation in corpora lutea counts, but not for both. We have no means of discriminating between these two possibilities on the basis of the present data.

(iv) Miscounting of corpora lutea. Perhaps the most likely reason for the reduced variation between sides in corpora lutea counts is miscounting. Crowded corpora lutea in a mouse ovary are not easy to distinguish, and close contiguity may often lead to two corpora lutea being counted as one. This would tend to reduce the counts for ovaries that shed a large number of eggs, and so diminish the variation between sides. This explanation is supported by the observation that the variation between sides declined as the total number of corpora lutea went up, though the decline was not statistically significant. On the other hand, miscounting on a scale sufficient to affect the variation between sides would be expected also to affect the mean. But a comparison of the means of corpora lutea counts and egg counts made on the same strains gives no evidence of miscounting (see Table 3). Although counts of corpora lutea as estimates of the numbers of eggs shed have often been criticized, the close correspondence between the corpora lutea counts and the egg counts in our data show that they can be reasonably reliable. Similar agreement was also reported by Falconer & Roberts (1960).

It is possible that several of the foregoing causes of increased or reduced variation were operating to produce the variation actually observed. The most likely cause of the reduced variation between sides in corpora lutea counts was miscounting in crowded ovaries. The increased variation between sides in egg counts was probably due to a difference in the blood supply to the two ovaries. After superovulation, however, there was little or no excess variation between sides in egg counts, though the difference in this respect from natural ovulation was not statistically significant. If the reduced variation between sides after superovulation was real, it could have resulted from the techniques employed: for example, the PMS and HCG injected intraperitoneally might reach the follicles partly through the surface of the ovary. This might then reduce any effect of a difference in the blood supplies to the two ovaries.

The striking effect that the dose of PMS had on the variation of the total number of eggs shed is puzzling. With doses of $\frac{1}{4}$ and $\frac{1}{2}$ i.u., the mean was similar to that of natural ovulation but the variance was considerably greater. With higher doses the variance was very much greater than would be expected from the extrapolation of data from natural ovulation. Thus the induction of superovulation by PMS introduced an additional cause of variation not present in naturally ovulating mice. What this cause of variation might be can only be surmised. It might be related to the stage of the oestrous cycle at which the PMS was injected, though there was no evidence that this had any effect (Fowler

& Edwards, unpublished). It might be due to the suppression of ovulation by excessive amounts of PMS in some mice through premature luteinization of follicles (Noble, Rowlands, Warwick & Williams, 1939; Fowler & Edwards, 1960). Or, it might depend on the rate of elimination of hormone from the circulation, though it would then have to be supposed that the differences between mice in their rates of elimination were more marked at high concentrations of PMS than at low.

One consequence of the greater variation after superovulation is that it produces a positive correlation between sides in place of the negative correlation found after natural ovulation. This change of sign need have no physiological implications other than those connected with the difference in variance. The correlations found are fully accounted for by the random distribution of eggs between the left and right ovaries together with the differences of the mean and variance of the total number of eggs shed.

ACKNOWLEDGMENTS

We are much indebted to Miss Mary Wheeler for carrying out the bulk of the statistical computations on which this work rests, and to Mr E. D. Roberts for drawing the figures.

REFERENCES

ADAMS, C. E. (1959) Studies on prenatal mortality in the rabbit, *Oryctolagus cuniculus*: the amount and distribution of loss before and after implantation. *J. Endocrin.* **19,** 325.

ALLEN, P., BRAMBELL, F. W. R. & MILLS, I. H. (1947) Studies on sterility and prenatal mortality in wild rabbits. I. The reliability of estimates of prenatal mortality based on counts of corpora lutea, implantation sites, and embryos. *J. exp. Biol.* **23,** 312.

BOWMAN, J. C. & ROBERTS, R. C. (1958) Embryonic mortality in relation to ovulation rate in the house mouse. *J. exp. Biol.* **35,** 138.

BRAMBELL, F. W. R. (1935) Reproduction in the common shrew (*Sorex araneus* Linnaeus). I. The oestrous cycle of the female. *Phil. Trans.* B, **225,** 1.

BRAMBELL, F. W. R. (1956) *Ovarian changes.* Marshall's Physiology of Reproduction, chap. 5. Ed. A. S. Parkes. Longmans, Green & Co., London.

BRAMBELL, F. W. R. & HALL, K. (1937) Reproduction of the lesser shrew (*Sorex minutus* Linnaeus). *Proc. zool. Soc. Lond.* p. 957.

BRAMBELL, F. W. R. & ROWLANDS, I. W. (1936) Reproduction of the bank vole (*Evotomys glareolus* Schreber). I. The oestrous cycle of the female. *Phil. Trans.* B, **226,** 71.

DANFORTH, C. H. & DE ABERLE, S. B. (1928) The functional interrelation of the ovaries as indicated by the distribution of foetuses in the mouse uteri. *Amer. J. Anat.* **41,** 65.

ECKSTEIN, P. & MCKEOWN, T. (1955) Effect of transection of one horn of the guinea-pig's uterus on foetal growth in the other horn. *J. Endocrin.* **12,** 97.

ECKSTEIN, P. & ZUCKERMAN, S. (1956) *Morphology of the reproductive tract.* Marshall's Physiology of Reproduction, chap. 2. Ed. A. S. Parkes. Longmans, Green & Co., London.

EDWARDS, R. G. & FOWLER, R. E. (1960) Superovulation treatment of adult mice: their subsequent natural fertility and response to further treatment. *J. Endocrin.* **21,** 147.

FALCONER, D. S. (1953) Selection for large and small size in mice. *J. Genet.* **51,** 470.

FALCONER, D. S. (1955) Patterns of response in selection experiments with mice. *Cold Spr. Harb. Symp. quant. Biol.* **20,** 178.

FALCONER, D. S. (1960a) Selection of mice for growth on high and low planes of nutrition. *Genet. Res.* **1,** 91.

FALCONER, D. S. (1960b) The genetics of litter size in mice. *J. cell. comp. Physiol.* **56,** Suppl. 1, p. 153.

FALCONER, D. S. & ROBERTS, R. C. (1960) Effect of inbreeding on ovulation rate and foetal mortality in mice. *Genet. Res.* **1,** 422.

FOWLER, R. E. & EDWARDS, R. G. (1957) Induction of superovulation and pregnancy in mature mice by gonadotrophins. *J. Endocrin.* **15,** 374.

FOWLER, R. E. & EDWARDS, R. G. (1960) The fertility of mice selected for large or small body size. *Genet. Res.* **1**, 393.

HOLLANDER, W. F. & STRONG, L. C. (1950) Intra-uterine mortality and placental fusions in the mouse. *J. exp. Zool.* **115**, 131.

JONES, E. C. & KROHN, P. L. (1961) The relationships between age, numbers of oocytes and fertility in virgin and multiparous mice. *J. Endocrin.* **21**, 469.

KENT, H. A. (1960) Polyovular follicles and multinucleate ova in the ovaries of young mice. *Anat. Rec.* **137**, 521.

LIPSCHÜTZ, A. (1928) New developments in ovarian dynamics and the law of follicular constancy. *J. exp. Biol.* **5**, 283.

MANDL, A. M., ZUCKERMAN, S. & PATTERSON, H. D. (1952) The number of oocytes in ovarian fragments after compensatory hypertrophy. *J. Endocrin.* **4**, 347.

NOBLE, R. L., ROWLANDS, I. W., WARWICK, M. H. & WILLIAMS, P. C. (1939) Comparative effects of certain gonadotrophic extracts on the ovaries of normal and hypophysectomized rats. *J. Endocrin.* **1**, 22.

ROBERTSON, A. (1951) The analysis of heterogeneity in the binomial distribution. *Ann. Eugen. Lond.* **16**, 1.

Tables of the Binomial Probability Distribution (1950) National Bureau of Standards, Washington.

PAPER 4

Embryonic mortality in relation to ovulation rate in the house mouse.

J. Exp. Biol. 35, 138-143. 1958.

by

J.C. BOWMAN & R.C. ROBERTS

REPRINTED FROM THE JOURNAL OF EXPERIMENTAL BIOLOGY,
VOL. 35, NO. 1, PP. 138–143, MARCH 1958.
 PRINTED IN GREAT BRITAIN

EMBRYONIC MORTALITY IN RELATION TO OVULATION RATE IN THE HOUSE MOUSE

BY J. C. BOWMAN
Institute of Animal Genetics, Edinburgh
AND R. C. ROBERTS
A.R.C. Animal Breeding Research Organization, Edinburgh 9

(*Received* 21 *September* 1957)

INTRODUCTION

An inverse relationship between the numbers of foetuses in the two horns of the uterus in the house mouse during the last week of gestation has been reported several times (Hollander & Strong, 1950; Runner, 1951). We recently found a significant negative correlation between the two ovaries in respect of the number of corpora lutea present at 18 days of gestation. In respect of implantation sites and live embryos at 18 days, however, the correlations were much less strongly negative and were non-significant. This reduction in the strength of the correlation points to a differential loss of eggs or of embryos, the horn receiving the greater number of eggs suffering a proportionately greater loss. We have accordingly re-investigated these correlations and looked for direct evidence of a differential loss. We have found an increasing proportional loss of eggs related to the number shed within a horn, and have shown that it occurs after fertilization but before implantation. The results are reported below.

RESULTS

Two series of observations were made. In the first, pregnant female mice of heterogeneous origin were dissected at 18 days of gestation. The numbers of corpora lutea were counted as a measure of the number of eggs shed from each ovary. The numbers of implantation sites and of live embryos in each uterine horn were also counted. The correlation between sides within mice for corpora lutea was $-0{\cdot}436$ ($P<0{\cdot}001$). In agreement with our previous findings, the numbers of implantation sites and live embryos showed lower negative correlations between sides than the number of eggs shed. The correlation between sides for implantation sites was $-0{\cdot}300$ ($P<0{\cdot}02$) and for live embryos $-0{\cdot}111$ (N.S.).

Further analysis revealed that the loss of eggs is affected by the number of eggs shed into the horn. Horns with a larger number of eggs suffer a proportionately greater loss. Thus, mice with an extreme distribution of eggs between sides will contribute largely to the negative covariance in corpora lutea counts, but as the two horns are affected by differential loss the negative correlation is greatly reduced by the time of implantation.

The distribution of loss of eggs and implanted embryos within uterine horns and

Table 1. *Distribution of amount and time of loss of eggs*

Corpora lutea per ovary	No. of ovaries examined	Total corpora lutea	Fraction of corpora lutea not accounted for by implants	Loss of implants up to 18 days of gestation		Fraction of corpora lutea not accounted for by live embryos at 18 days of gestation
				As a fraction of eggs shed	As a fraction of implants	
A. Within uterine horns						
1	0	0	0	0	0	0
2	6	12	0·083	0	0	0·083
3	10	30	0·066	0·066	0·071	0·133
4	24	96	0·135	0·063	0·072	0·198
5	26	130	0·085	0·092	0·101	0·177
6	28	168	0·137	0·083	0·097	0·220
7	20	140	0·171	0·029	0·034	0·200
8	9	72	0·278	0·042	0·056	0·319
9	7	63	0·254	0·222	0·298	0·476
10	2	20	0·150	0	0	0·150
Total	132	731	0·155	0·075	0·089	0·230
B. Within mice						
Corpora lutea per mouse	No. of mice examined					
8	6	48	0·125	0·104	0·119	0·229
9	12	108	0·074	0·074	0·080	0·148
10	6	60	0·133	0·117	0·135	0·250
11	15	165	0·158	0·060	0·072	0·218
12	11	132	0·106	0·053	0·059	0·159
13	10	130	0·169	0·062	0·074	0·231
14	3	42	0·167	0·119	0·143	0·286
15	2	30	0·633	0·133	0·364	0·767
16	1	16	0·188	0·062	0·077	0·250
Total	66	731	0·155	0·075	0·089	0·230

Table 2. χ^2 *analysis of loss of eggs*

Source of variation	D.F.	χ^2	M.S.	F	P
A. Within uterine horns					
(1) Eggs lost up to implantation					
Linear trend	1	13·166	13·166	10·206	0·02
Deviations from linear trend	7	9·028	1·290	8·354	0·01
Heterogeneity of loss within corpora lutea groups	123	193·885	1·576		
(2) Eggs not accounted for by live embryos at 18 days					
Linear trend	1	14·937	14·937	5·888	0·05
Deviations from linear trend	7	17·760	2·537	9·215	0·01
Heterogeneity of loss within corpora lutea groups	123	199·430	1·621		
B. Within mice					
(1) Eggs lost up to implantation					
Linear trend	1	18·938	18·938	2·571	N.S.
Deviations from linear trend	7	51·565	7·366		N.S.
(2) Eggs not accounted for by live embryos at 18 days					
Linear trend	1	13·501	13·501	1·836	N.S.
Deviations from linear trend	7	51·485	7·355		

within mice is shown in Table 1. The χ^2 analyses for linear trend of proportion lost at two stages of gestation are given in Table 2. Only the linear trends of loss within horns up to implantation and for total loss up to 18 days of gestation are significant. The loss within horns between implantation and 18 days of gestation is very irregular and does not seem to be related either to the number of eggs shed into the horn or to the number of eggs implanted. We conclude therefore that before implantation there is positive linear trend of loss of eggs with increasing numbers shed per horn. This conclusion leads us to expect to find a similar increasing proportional loss when the data are analysed on a within-mouse basis. Our data does suggest such a trend, though statistical analysis reveals this to be non-significant.

Table 3. *Fraction of eggs not fertilized in relation to number shed per ovary*

Eggs shed per ovary	No. of Fallopian tubes examined	Total no. of eggs	Fraction not fertilized
1	2	2	0·500
2	7	14	0·429
3	7	21	0·048
4	12	48	0·125
5	15	75	0·173
6	11	66	0·030
7	13	91	0·132
8	8	64	0·125
9	1	9	0·000
10	5	50	0·080
11	—	—	—
12	1	12	0·250
Total	82	452	0·126

The differential loss between horns up to implantation may be due either to a lack of fertilization or to a failure to implant, and a second series of observations was made to solve this problem. Females of heterogeneous origin, similar to those used in the first series of observations, were put singly with a male between 5.0 and 5.30 p.m. and examined the following day between 9.0 and 10.0 a.m. for vaginal plugs. Those which had mated were killed between 7.0 and 10.0 p.m. and the eggs in each Fallopian tube were extracted. From the findings of Snell, Fekete, Hummel & Law (1940), Snell, Hummel & Abelmann (1944) and Braden & Austin (1954) we considered that the majority of fertilized eggs would at that time be in the pronucleate stage. The eggs were examined by phase-contrast microscope according to the method described by Austin & Smiles (1948). The numbers of fertilized and non-fertilized eggs in each tube were counted. Judgement as to whether eggs were fertilized or not was based on the description by Austin (1951) of the formation of the pronuclei in the rat egg.

The correlation between ovaries in the numbers of eggs shed in a mouse was found to be $-0{\cdot}528$ (P 0·001), which is in good agreement with the similar correlation for corpora lutea counts. The number of eggs not fertilized per tube in relation to the total eggs shed per ovary are shown in Table 3. In marked contrast to the

implantation data, no regular trend of loss is apparent in this case. The data has also been analysed on a within-mouse basis and as expected no trend is shown. These results indicate that fertilization rate is not related to the number of eggs shed per ovary or per mouse and that it does not normally limit litter size in the mouse.

DISCUSSION

Our observations indicate that the loss of implanted eggs up to 18 days of gestation does not vary with the number of implantations in a horn, but that as the number of eggs shed into a uterine horn increases the probability of each individual egg implanting decreases. The fertilization rate is not related to the number of eggs in the horn, and therefore the factor or factors causing the variation in implantation rate must be operating on fertilized eggs or on pre-implantation embryos.

Our results and conclusions do not entirely agree with those of other workers. Danforth & de Aberle (1928) and McLaren & Michie (1956) found no correlation between the two horns for the number of implantations. As mentioned earlier other authors have reported significant negative correlations, and therefore we can only attribute the inconsistency of the results reported to heterogeneity between mice used at different laboratories.

The differential loss of eggs in our data could be explained if trans-uterine migration of eggs had occurred in many of our mice. Such migration is known to occur in rodents (Runner, 1951; Boyd & Hamilton, 1952; Young, 1953; McLaren & Michie, 1954), but these reports suggest that its frequency is very low. For this reason we have dismissed migration as an explanation of our results.

Previous work and ideas as to the causes of pre-implantational loss have been reviewed by Hammond (1952), but it is impossible to decide from our present experimental evidence which if any of the causes are applicable to our findings. Some useful conclusions may be made, however, by comparing our results with those of McLaren & Michie (1956).

From the results of an experiment, in which they transferred varying numbers of $3\frac{1}{2}$-day-old blastocysts from donor mice to normally mated recipient mice $2\frac{1}{2}$ days pregnant, they concluded that 'although we have found no limit to the number of eggs which can implant in a single uterine horn, we are beginning to approach a limit to the number of implantations which a single horn can keep alive'. In the data reported here the post-implantational loss was irregular and not proportionately related to the number of implants in the horn. The reason for this apparent difference is fairly easily found. McLaren & Michie (1956) consider that in their material the limit to the number of implantations which remain alive to 16 days of gestation may possibly be due to 'insufficiency of corpora lutea to supply the progesterone requirements of the excessive number of implantations'. In all animals included in our data, presented here, there were at least as many corpora lutea in the ovary as implantations in the horn to which it corresponded, and consequently the progesterone supply is much less likely to have been insufficient. It is possible, however, that the total number of implants per mouse surviving to birth is limited

by the level of some substance circulating in the maternal blood supply—a hypothesis favoured by Runner (1951) and by Hammond (1952).

McLaren & Michie (1956) found in their experiment that the number of successful implantations, from donor and recipient sources combined, rises linearly by increments of 0·2 for each additional egg injected. Again, this conclusion seems to conflict with our results of an increasing proportional loss of eggs up to implantation. However, the discrepancy may not be so serious as it first appears. If indeed there is no increase in the fractional loss in McLaren & Michie's data, then it seems highly probable that the increased proportionate mortality in our material occurred very soon after fertilization, in fact between fertilization itself and the stage at which the blastocysts were removed by McLaren & Michie for transplantation.

If this suggestion is correct we have a more accurate estimate of the time interval during which the differential loss of eggs takes place, and this knowledge might prove useful in further elucidation of causes of pre-implantational loss.

SUMMARY

1. Two series of observations were made to determine the time and amount of loss of eggs in relation to the number shed per ovary and per mouse.

2. The correlations between sides within mice for eggs shed was −0·528, for corpora lutea counts −0·436, for implantations −0·300, and for live embryos −0·111.

3. A positive linear trend of loss of fertilized eggs with increasing numbers of eggs per uterine horn has been shown to occur before implantation.

4. Possible causative mechanisms for the loss are discussed in relation to observations on embryonic mortality previously reported by other workers.

We wish to express our gratitude to Prof. C. H. Waddington for laboratory facilities, to Dr D. S. Falconer for much helpful criticism and advice and to Dr B. Woolf for statistical advice. J. C. Bowman gratefully acknowledges financial support from the Agricultural Research Council.

REFERENCES

AUSTIN, C. R. (1951). The formation, growth, and conjugation of the pronuclei in the rat egg. *J. R. Micr. Soc.* **71**, 295–306.

AUSTIN, C. R. & SMILES, J. (1948). Phase-contrast microscopy in the study of fertilization and early development of the rat egg. *J. R. Micr. Soc.* **68**, 13–19.

BOYD, J. D. & HAMILTON, W. J. (1952). Marshall's *Physiology of Reproduction*, Vol. II, chap. 14. London: Longmans.

BRADEN, A. W. H. & AUSTIN, C. R. (1954). Fertilisation of the mouse egg and the effect of delayed coitus and of hot shock treatment. *Aust. J. Biol. Sci.* **7**, 552–65.

DANFORTH, C. H. & DE ABERLE, S. B. (1928). The functional interrelation of the ovaries as indicated by the distribution of foetuses in mouse uteri. *Amer. J. Anat.* **41**, 65–74.

HAMMOND, J. (1952). Marshall's *Physiology of Reproduction*, vol. II, chap. 21. London: Longmans.

HOLLANDER, W. F. & STRONG, L. C. (1950). Intra-uterine mortality and placental fusions in the mouse. *J. Exp. Zool.* **115**, 131–47.

MCLAREN, A. & MICHIE, D. (1954). Transmigration of unborn mice. *Nature, Lond.*, **174**, 844.

MCLAREN, A. & MICHIE, D. (1956). Studies on the transfer of fertilised mouse eggs to uterine foster mothers. *J. Exp. Biol.* **33**, 394–416.

RUNNER, M. N. (1951). Differentiation of intrinsic and maternal factors governing intra-uterine survival of mammalian young. *J. Exp. Zool.* **116**, 1–20.

SNELL, G. D., FEKETE, E., HUMMEL, K. P. & LAW, L. W. (1940). Relation of mating ovulation and the oestrus smear in the house mouse to time of day. *Anat. Rec.* **76**, 39–54.

SNELL, D., HUMMEL, K. P. & ABELMANN, W. H. (1944). A technique for the artificial insemination of mice. *Anat. Rec.* **90**, 243–53.

YOUNG, A. (1953). Transuterine ('Internal') migration of fertilised ova in the albino rat. *Proc. Roy. Soc.* Edinb. B, **65**, 106–16.

PAPER 5

The limits to artificial selection for body weight in the mouse. I. The limits attained in earlier experiments.

Genet. Res. Camb. 8, 347-360. 1966

by

R.C. ROBERTS

Genet. Res., Camb. (1966), **8**, *pp.* 347–360
With 3 text-figures
Printed in Great Britain

The limits to artificial selection for body weight in the mouse

I. THE LIMITS ATTAINED IN EARLIER EXPERIMENTS

By R. C. ROBERTS

A.R.C. Unit of Animal Genetics, Institute of Animal Genetics, Edinburgh, 9

(*Received* 21 *June* 1966)

1. INTRODUCTION

The expected pattern of response to artificial selection is well known—progress is made at an ever diminishing rate as the limit is approached asymptotically. Though deviations from this general form are frequently encountered in practice, as discussed by Falconer (1955), ultimately a stage is reached after which no further progress is made. This limit to selection will inevitably be met when all the alleles affecting the trait have been fixed in the population; in biometrical terms, the genetic variance will then have been exhausted. But the limit may be reached well before the point when the genetic variance is exhausted, and despite the fact that some loci are not fixed, selection may fail to change the mean value of the population any further. Such a contingency may arise if the selection favours individuals that are heterozygous at some loci, of if natural selection opposes the direction of the artificial selection.

In view of such uncertainties about the nature of the limit to selection, predictions about the length of time taken to reach the limit, and its ultimate level, become hazardous. Reviewing some experimental evidence from mice and *Drosophila*, Falconer (1960*a*) suggests that the response may be expected to continue for some twenty or thirty generations, producing a total divergence between strains selected for high and low expressions of the trait of the order of fifteen to thirty times the additive genetic standard deviation in the initial population, or ten to twenty times the phenotypic standard deviation. In a theoretical treatment of the subject, Robertson (1960) formulates his conclusions in terms of the effective population size, N. Robertson confirms Dempster's (1955) derivation that the total advance should equal $2N$ times the gain in the first generation, provided that the rate of fixation is low and provided also that the genes act additively. If dominance is involved, the total advance may be well in excess of this amount. Robertson shows also that half of the total gain should be achieved in not more than $1{\cdot}4N$ generations for genes that act additively, though the figure may rise to $2N$ generations for rare recessives. If the half-life of the selection process falls short of $1{\cdot}4N$ generations, Robertson suggests that the majority of alleles favourable to the direction of the selection will have been fixed in the population.

An important concept involved in a discussion of selection limits is this chance fixation of some unfavourable alleles in a selected line even though selection is directed against them. The probability that this may occur will obviously depend on the population size, and also on the selective advantage of the gene, or the intensity of selection in a quantitative situation. Kimura's (1957) treatment of chance fixation is extended by Robertson to show that the expected limit to selection based on individual measurements is a function only of the product Ni (where i is the intensity of selection, measured as the selection differential in phenotypic standard deviation units). As Ni increases, the probability diminishes that the less favourable allele at a locus is fixed during the course of selection.

A limitation on Robertson's theoretical treatment is that it is developed entirely in terms of the exhaustion of additive genetic variance. The study of selection limits is therefore still largely confined to the experimental investigation of particular cases. The present series of papers will report some long-term experiments on the limits to artificial selection for body weight in the mouse. This first paper reviews the limits attained in earlier selection programmes in this laboratory. Later papers will examine more closely the genetic nature of the limits, and will describe methods whereby further progress might be made.

2. MATERIAL AVAILABLE FOR STUDY

Seven selected strains of mice—four large and three small ones—were available for study in this laboratory. As far as can be judged, each strain had been selected to its limit for body weight, either high or low as the case may be. The designation of these strains, the number of generations of selection they had undergone prior to this study, and references to their original sources are all shown in Table 1. Briefly,

Table 1. *Strains selected to the limit for body weight*

Line	Generation reached prior to present study	Character selected	Reference
RCL	36	High 6-week weight	Falconer & King, 1953
NF	52	High 6-week weight	Falconer, 1953
CFL	31	High growth, 3–6 weeks	Falconer, 1960*b*
CRL	31	High growth, 3–6 weeks	Falconer, 1960*b*
MS	38	Low 6-week weight	MacArthur, 1949; King, 1950
NS	42	Low 6-week weight	Falconer, 1953
CFS	31	Low growth, 3–6 weeks	Falconer, 1960*b*

the origin of the various strains was as follows. *RCL* stemmed originally from a cross between Goodale's (1938, 1941) and MacArthur's (1944, 1949) large strains. The *NF* and *NS* strains both derived from a four-way cross of inbred lines. *CFL* and *CFS* were selected from a heterogeneous outbred base population, but one which contained *RCL* and had also some overlap with the *N* strains; *CRL* had an identical origin but was selected on a low plane of nutrition. *MS* stands for 'MacArthur's

Small', but is a slight misnomer. Dr MacArthur supplied nine males to this laboratory in 1948. These were crossed with females of three inbred strains. Some of the original males were available for three further backcrosses, though these matings were supplemented with some intercrosses. The result was a population 87% of whose genes derived from the original MacArthur strain, which formed a base population for further selection for small size.

In every case, the selection was within litters, to avoid some of the complications due to maternal effects in the interpretation of the results. The character selected was either the body weight of the mouse at 6 weeks of age or else the growth between 3 and 6 weeks. These two characters are scarcely distinguishable in terms of the ranking of the mice on the two measurements (Falconer, 1955), which enables us to discuss both sets of experiments within the same framework. The limits reached are examined empirically and in terms of Robertson's theory, with its extension by Hill (1965) and Hill and Robertson (1966).

3. RESULTS AND DISCUSSION

(i). *Empirical observations*

A summary of the responses to selection of the seven strains available in the laboratory is given in Figs. 1, 2 and 3. The mean weights are plotted against the number of generations of selection; in the present context, this is the most meaningful way to examine the results. The present analysis is confined to the limits ultimately reached. We are not concerned here with the patterns of the response nor with other features discussed in the original publications. However, some points that have arisen since those publications are relevant to the present discussion. Figure 1 is straightforward, but Fig. 2 presents a complication. There was a decline in weight of the *CFL* line between generation 19 and generation 27, and no ready explanation is available. It is too great to dismiss as an accident of sampling, and as the other two selected lines in the same figure were mated contemporaneously with *CFL*, a general environmental trend cannot be invoked. For whatever reason, the outcome was that the *CFL* ultimately reached a level not much above its origin. However, the decline in the *CFL* line assumes less significance when compared to the precipitous fall in weight of the *RCL* line, shown in Fig. 3. Between generations 19 and 24, the mean weight dropped by no less than 16 g., despite continued selection for large size. Although there was some recovery in later generations, the *RCL* line never again achieved its previous high weights, and provides a second instance of a selected line ending up more or less where it began. Newman (1960) investigated the rise and fall of the *RCL* line in some detail. He carefully excluded the possibility of an accidental outcross to a smaller line and, by comparing expected and realized selection differentials, he failed to establish that there was any natural selection against large size over this period. In fact, the magnitude of the decline is not amenable to any reasonable genetic interpretation, and Newman was forced to postulate the imposition of some environmental stress, possibly an unidentified pathogen, that was highly specific to the *RCL* line. Nevertheless, whatever the

cause, a genetic change of an unfavourable kind was brought about, otherwise the line should ultimately recover its previous level. From Fig. 3, it can be seen that the supposition of eventual recovery would, at best, invite scepticism.

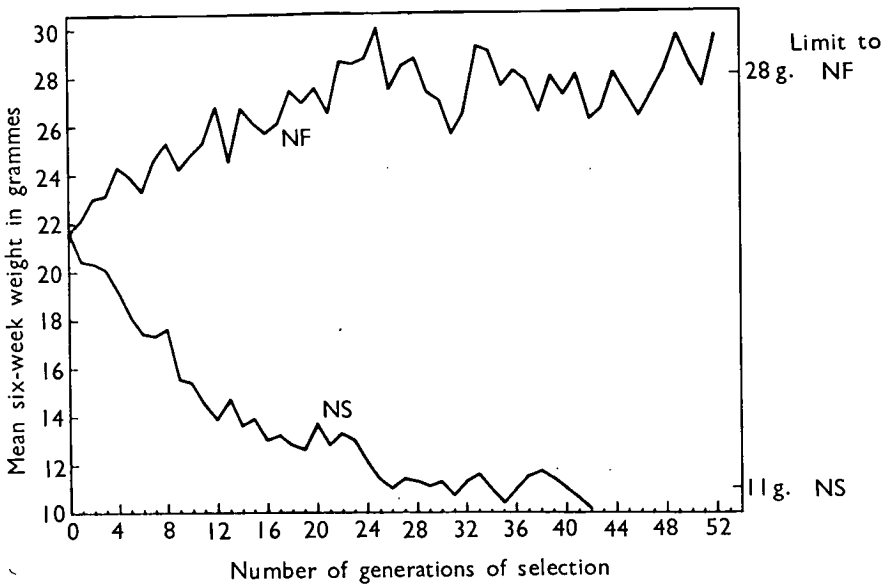


Fig. 1. Responses to selection for body weight. The limits attained in selected lines first reported by Falconer (1953).

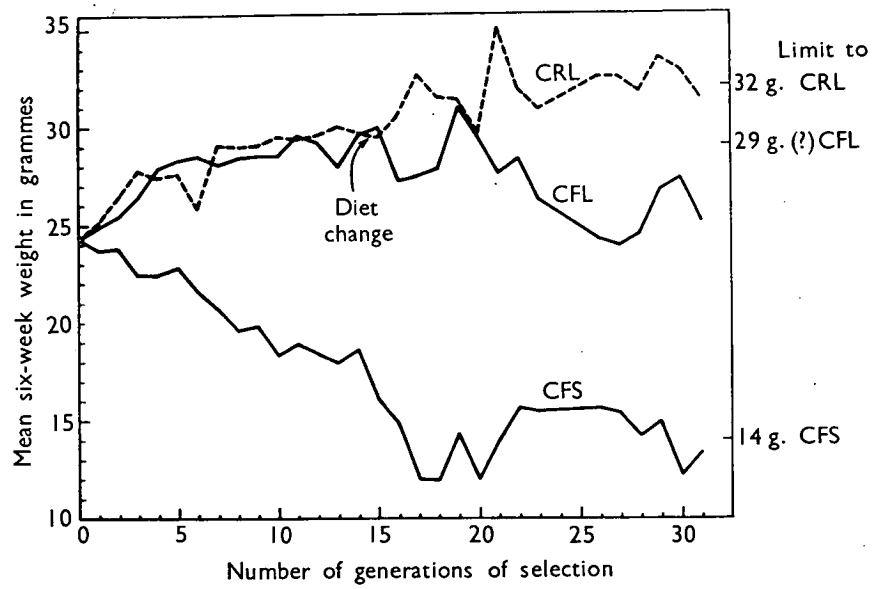


Fig. 2. Responses to selection for growth. The limits attained in selected lines first reported by Falconer (1960*b*). *CRL* was selected on a restricted diet for fourteen generations, but the weights shown were for animals measured on a normal diet. Only the criterion of selection changed at the point marked 'diet change'.

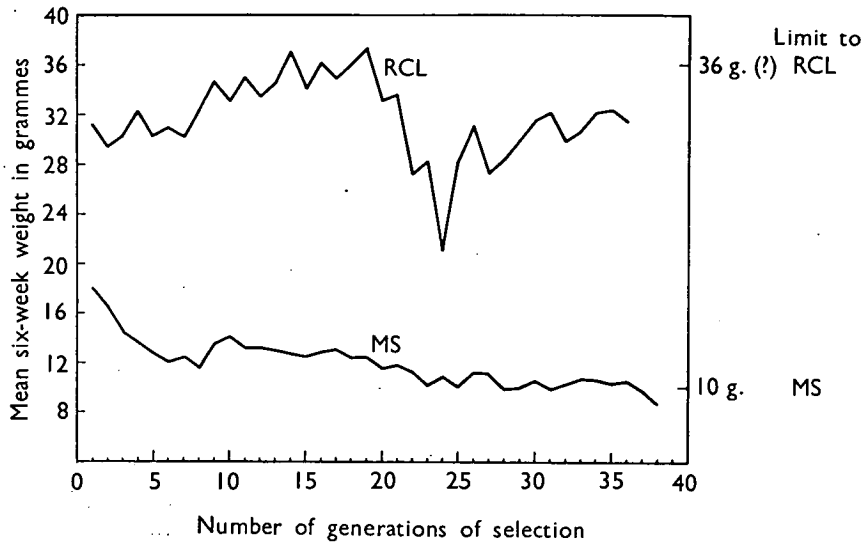


Fig. 3. Responses to selection for body weight. Upper graph: the limit attained in selected line first reported by Falconer & King (1953). Lower graph: the limit attained in selected line first reported by MacArthur (1944, 1949) and reconstructed by King (1950).

The picture that emerges from all this is that, at the limit, lines of mice selected for large size tend to be rather unstable. At the very least, we cannot regard the attainment of a steady state at the limit as an inviolable rule. Even the *NF* line, which shows the clearest pattern, is inclined to oscillate rather violently between higher and lower weights, although over a longer period no discernible trend is apparent. This is also a feature of one of the small lines (*CFS*). For this reason, it becomes extremely difficult to decide what mean weight we are prepared to regard as 'the limit', and quite impossible to decide at what exact point in time this limit was reached. As a rough guide, some weights have been marked on the right-hand sides of Figs. 1, 2 and 3, showing the approximate limits reached. The weights shown were derived quite subjectively, by averaging to the nearest gramme the mean weight over the period during which the line concerned was at its highest or lowest, as appropriate, and showed no obvious trend. It is fortunate perhaps that for present purposes, any more precise estimates would have served no purpose. The question marks after the limits shown for *RCL* and *CFL* are there for reasons that are all too obvious; the limits marked correspond to what looked like the limit before these lines declined.

In similar fashion, the time taken to reach the limit has been taken rather arbitrarily as the generation that first exceeded the level of the limit. If we think of the hypothetical smooth curve approaching an asymptote, it can be appreciated that accidents of sampling will tend to make the criterion an underestimate of the number of generations required. However, in the absence of a clear alternative, we shall accept this estimate, bearing in mind that it is probably biased downward.

The level of the limit in absolute terms is less interesting than the magnitude of

the response in terms of the variance in the base population before any selection was practised. The most informative way of looking at the response is in 'standard units', i.e. as multiples of the original standard deviation. Falconer (1955) gives the requisite information for the *NF* and *NS* lines; the phenotypic standard deviation was 1·9 g., while the additive genetic standard deviation was 0·9 g. The corresponding figures for the *C* stocks were 2·3 g. and 1·3 g.; these values were calculated from data on the base population kindly provided for me by Dr Falconer.

The results derived by these admittedly somewhat crude methods are presented in Table 2. The *RCL* and *MS* lines, by the nature of their origin, mentioned earlier, represent a situation totally different from the other five lines. Their mean levels

Table 2. *Limits reached by selected lines of mice*

Line	Limit in grammes	Generations to reach limit	Response		
			Grammes	$/\sigma_P$	$/\sigma_A$
RCL	36	14	4·8	—	—
NF	28	22	6·4	3·4	7·1
CFL	29	11	4·7	2·0	3·6
CRL	32	17	7·7	3·3	5·9
MS	10	28	8·0	—	—
NS	11	26	10·6	5·6	11·8
CFS	14	17	10·3	4·5	7·9

The last two columns evaluate the response as multiples of σ_P and σ_A respectively, where σ_P is the phenotypic standard deviation in the base population, and σ_A is the additive genetic standard deviation.

are presented for comparison with the other lines with a shorter history of selection, but beyond that they cannot be discussed in the same context. The apparent response of the *MS* is false in any event; most of it occurred during the first few generations and represents the repeated backcrosses after an outcross as mentioned previously.

Table 2 permits some empirical statements about the limit to artificial selection for body weight in the mouse at 6 weeks of age. It is emphasized that this is a well-defined character and that the experiments were all conducted in the same laboratory over much the same period of time. The outcome was that superficially, different experiments were in broad qualitative agreement with each other. Some large mice were developed that had mean weights in the region of 30 g., while the small mice ceased to respond around 12 g., give or take a gramme or two at both levels. Yet, when these separate lines are examined more closely in terms of the limits reached, some important differences emerge. Firstly, the response may continue for anything, it seems, between ten and thirty generations. On a temporal scale, this represents for the mouse a range from, at best, 2 years to, at worst, 8 years. Translating the result to domestic livestock, where the generation interval may well exceed 2 years, this range assumes far greater importance. It becomes desirable, therefore, to scan the base populations for reliable correlates of the duration of the

response, and to evaluate the effects of such correlates on the limit ultimately reached. Unfortunately, excluding the irrelevant cases of *RCL* and *MS*, the lines discussed here were derived from two base populations only, and correlations based on only two points do not engender much faith. But for what they are worth, the following observations can be made from Table 3, which derives largely from the

Table 3. *Duration of response in relation to variances in base populations*

Lines	Generations to reach limits	Base population			Response	
		σ_P	σ_A	h^2	$/\sigma_P$	$/\sigma_A$
C	15	2·3	1·3	0·31	3·3	5·8
N	24	1·9	0·9	0·22	4·5	9·5

σ_P and σ_A are defined in legend of Table 2. h^2 is the heritability $=\sigma_A^2/\sigma_P^2$.

arithmetical means of some quantities presented in Table 2, and the information given previously about the base populations. The *C* lines reached the limit in less time than the *N* lines, and the base population of the *C* lines showed larger variances and a higher heritability. Since such a population would be preferred for selection purposes anyway, there is no incompatibility of objectives on this score. However, by virtue of the longer time taken to reach the limit, the final response of the *N* lines was just as impressive as that of the *C* lines, suggesting that their lower genetic variance had somehow been utilized more effectively. The material on which these observations are based is too tenuous to warrant further speculation, especially as other variables affect the limit attained. But it may serve to focus attention on the kind of information that is required.

A final point on the duration of the response is that no differences appear between large and small mice in this respect. The differences that were observed seem to be associated entirely with features of the base populations.

In terms of the variances in the base populations, it appears from Table 2 that the final response may amount to between two and six times the phenotypic standard deviation, and anything between three and twelve times the additive genetic standard deviation. These values were calculated for the response in one direction only. For the total divergence in two-day selection, values for corresponding high and low lines should be added together. When this is done, it puts the *C* lines, especially, slightly lower than the bottom of the range suggested by Falconer, quoted earlier.

The results obtained from the selection experiments discussed in this section must now be examined against the theoretical considerations outlined earlier.

(ii) *Theoretical considerations*

The theory of limits (Robertson, 1960; Hill, 1965; Hill & Robertson, 1966) outlined earlier frames its conclusions in terms of the effective size (N) of the population. We must therefore estimate the effective sizes of the populations under discussion.

The number of matings used to propagate the stocks during selection was not constant from generation to generation. Some of the variation was deliberate, as different numbers of mice were required for different phases of the experiments. Most of the variation, however, was attributable to some sterility, which is a common feature of all selected stocks. The procedure under such circumstances is quite straightforward. The effective number is given by the harmonic mean of the number of individuals that contributed to the succeeding generation. The results for the seven lines are shown in Table 4.

Table 4. *Half-life of selection responses*

Line	Effective number N	Half-life (generations)	Values of $Ni\alpha$ $p=0{\cdot}5$	0·25	0·1
RCL	19·0	$9=0{\cdot}47N$	—	—	—
NF	14·5	$8=0{\cdot}55N$	6	8	(10)
CFL	15·8	$4=0{\cdot}25N$	10	14	20
CRL	16·8	$7=0{\cdot}42N$	7	9	12
MS	19·5	$4=0{\cdot}21N$	—	—	—
NS	14·6	$9=0{\cdot}62N$	4	5	(8)
CFS	18·8	$10=0{\cdot}53N$	5	7	10

p is the frequency in the base populations of genes favourable to the direction of the selection. Values of 0·1 were not possible for the *NF* and *NS* lines, from the method of their construction.

It was mentioned earlier that the method of selection adopted was in all cases within families. It is well known that in idealized populations, this practice ought to double the effective number; each family contributes two individuals as parents for the next generation, which reduces to zero the variance between families in their contribution. However, mouse stocks always show some sterility, and to obtain the requisite number of matings, some families (and especially the larger ones) will contribute more than two individuals as parents for the next generation. It becomes imperative then to determine how these complications should be accommodated to estimate the effective number. The proper approach under such circumstances is to compute from pedigrees the inbreeding coefficient accumulated during the selection. If the inbreeding coefficient after t generations is F_t, then the formula

$$F_t = 1-\left(1-\frac{1}{2N}\right)^t$$

can be solved to give the effective number, N. I am indebted to Dr D. S. Falconer for kindly providing me with some inbreeding coefficients he had calculated for the *NF* and *NS* stocks. The effective numbers, as established by this accurate method, compare with the estimates from the harmonic mean over the same period as follows:

Stock	Generations	Effective number from inbreeding coefficients	Harmonic mean
NF	26	14·9	13·1
NS	22	14·3	13·6

It is seen that the harmonic mean provides an estimate that is only slightly lower than the accurate calculation, whereas in idealized populations one should be half the other. This does not imply that the within-family method of selection did not increase the effective number over what it would have been with, say, mass selection. Without regard to the representation of as many families as possible, variation in fertility and viability leads to an effective number much lower than the supposed number of parents.

As the N stocks did not appear to differ much from the others with respect to fertility and ease of maintenance, we shall accept the harmonic mean of the number of parents as being a sufficiently accurate estimate of the effective number for all the stocks. It is possibly a slight underestimate of the true value, but any error that may be involved is not sufficient to affect grossly any conclusions that we may draw.

We shall now examine the half-life of the selection response in terms of the effective population size. The half-life was estimated in a manner identical to that explained in connexion with the total response. In this case, the half-life was taken as the generation whose mean first exceeded one-half of the total response. Again, this will tend to underestimate the true value. The results, tabulated in Table 4, reveal that half of the response was obtained in most cases by about $\frac{1}{2}N$ generations, whereas the value expected when the chance of fixing an unfavourable allele is not high varies at most from N to $2N$ generations, as shown by Robertson (1960). The implication of this low value of half-life, in the context of a study of selection limits, is that all of the alleles favourable to the direction of the selection should have been fixed. Should it turn out that a less favourable allele has been fixed, then the disparity between the value of $\frac{1}{2}N$ and the range quoted by Robertson is such that we may safely infer that some process other than fixation is operative in the determination of the limit reached.

The values obtained for the half-life of the selection process lead directly to two other estimates that are of some consequence in quantitative genetics. The first reflects the order of magnitude of the effect of the individual genes involved in the response to the selection. The second provides some estimate of the number of 'loci' or effective factors which are concerned in the process. This number of course estimates only those loci which happen to be segregating in that particular population. Though these estimates are by their nature imprecise, they cover an area where but little knowledge is available, especially for mammals.

The procedure for estimating the gene effects and the number of loci is most easily derived as follows. It can be shown (Robertson, 1960, as developed by Hill, 1965) that a half-life of a given magnitude corresponds to a limited range of values of $Ni\alpha$. N, the effective population size, has been discussed already; i is the intensity of selection, and tabulated values in terms of the proportion of animals selected are widely available; α is the average proportionate effect of the genes:

$$\alpha = \frac{a}{\sigma_P}$$

where a is defined as the difference in value between the two homozygotes, and σ_P is the phenotypic standard deviation.

Now, the exact value of $Ni\alpha$ corresponding to a certain half-life depends somewhat on the gene frequencies in the base population. Some graphs are provided by Hill & Robertson (1966), and by interpolation, values corresponding to the appropriate half-life and specified gene frequencies may be obtained. Such values, for gene frequencies of 0·5, 0·25 and 0·1 in the base populations, are entered in Table 4. The *RCL* and *MS* lines are ignored since their previous history excludes them from being subjected to the present treatment. The values for the *NF* and *NS* lines corresponding to a gene frequency of 0·1 are entered in parentheses, since frequencies lower than 0·25 were impossible in this stock from the method of its construction.

Thus, having estimated $Ni\alpha$, we may now derive α, since N and i are observable quantities. N is given in Table 4, and i for the selected lines described here was always close to 1·0. This value was ascribed to all lines, being quite accurate enough for present purposes. However, the value of α so obtained must be adjusted to allow for the fact that the selection was, in all cases, based on deviations from the means of full-sib families. The selection therefore operated on only half of the additive genetic variance in the population, and the corresponding phenotypic variance is that within families (σ_w^2). The definition of α must therefore be modified appropriately:

$$\alpha = \frac{a}{2\sigma_w}$$

Since we still want to derive the proportionate effect of the genes on a population basis, let

$$k = \frac{\sigma_w}{\sigma_P}$$

Then, the proportionate effect of the genes (a/σ_P) is given by:

$$\frac{a}{\sigma_P} = 2k\alpha$$

Values of k were calculated for the base populations from data kindly supplied by Dr D. S. Falconer. These were employed to estimate the proportionate effects.

Now, to estimate the number of loci involved in the response, we need to consider the within-family heritabilities (h_w^2), published for the N stocks by Falconer (1955) and for the C stocks by Falconer (1960*b*). Values for the high and low lines were averaged, and the average taken to apply to the base population. Each locus, in the terms outlined above and in a within-family selection programme, contributes $\frac{1}{4}a^2p(1-p)$ to the additive genetic variance, where p is the gene frequency. If we make the assumption that each of the loci involved contributes equally to the genetic variance, then

$$h_w^2 = \frac{na^2p(1-p)}{4\sigma_w^2}$$

where n is the number of loci contributing to the response. By rearranging the expression derived, we obtain

$$n = \frac{4\sigma_w^2 h_w^2}{a^2 p(1-p)}$$

$$= \frac{h_w^2}{\alpha^2 p(1-p)}$$

Since α and h_w^2 have already been determined, this enables us to estimate the number of loci by substituting various values for the initial gene frequency.

The estimates of the average proportionate effects of the genes and the number of loci concerned are shown in Table 5, for the five lines to which the procedure was

Table 5. *Proportionate effects of genes and number of loci*

		Proportionate effect (α)			Number of loci		
Line	h_w^2	$p=0{\cdot}5$	0·25	0·1	$p=0{\cdot}5$	0·25	0·1
NF	0·35	0·57	0·76	(0·95)	8	6	8
CFL	0·33	1·03	1·46	2·08	3	2	2
CRL	0·33	0·59	0·79	0·98	10	8	10
NS	0·35	0·37	0·47	(0·76)	19	16	13
CFS	0·33	0·44	0·61	0·87	18	13	13

h_w^2 is the realized heritability within litters.

applied. Over the range of gene frequencies considered, the estimated number of so-called loci does not very much, since $p(1-p)$ diminishes as α^2 increases. But above a gene frequency of 0·5, both would tend to diminish together, leading to successively lower values for the number of loci, though $Ni\alpha$ (and therefore α) does not alter much over this range.

The estimates shown in Table 5 are not given with any pretensions about their numerical accuracy. Rather, they serve as indicators of the order of magnitude of the effects with which we are dealing. By and large, however, the five lines have produced reasonably consistent answers. They seem to indicate that the average difference between the two homozygotes at a locus produces an effect usually in the region of a half to one phenotypic standard deviation, and that this corresponds to a total of up to twenty loci in the base population contributing to the response to selection. If some of the estimates of the number of loci appear to be low, it should be noted that any violation of the basic assumptions biases the estimate downwards. The fact that the lines selected for small size appear to have more loci contributing to the response does not arouse much curiosity. Directional dominance favours large size in the mouse. If selection is for the dominant genes, this leads to a shorter half-life, a higher value of $Ni\alpha$ and thus to a lower estimate of the number of genes, if other factors remain constant.

The estimates obtained of the proportionate effect of the genes and the number of loci involved perhaps serve three purposes. Firstly, they can be compared with some other meagre evidence on the same topic. For instance, Falconer (1960*a*) gives estimates derived by an alternative (though related) approach for some traits in both mice and *Drosophila*; his figures for 6-week weight in the mouse are of the same order of magnitude as the ones given here. Secondly, the estimates reveal no basic incompatibility between the parameters of the base populations and the responses actually obtained. And lastly, they lend some experimental support to the theoretical considerations developed by Robertson and by Hill.

4. CONCLUSIONS

This survey of previous selection experiments for body weight indicates to within a fairly narrow range the limits that can be expected, under the conditions of our laboratory, when selection is applied to a heterogeneous population. It seems that the upward response reaches its limit around 30 g. while the downward response ceases in the region of 12 g. or so, on average. The most extreme cases found were a high line limit of 32 g. (unless we invoke the transient glory of the *RCL* line before its mysterious decline) and a low line limit of 10 g. These figures set standards for further experimental attacks on the limits.

What is also of relevance in this context is that from theoretical considerations, we have been able to exclude almost completely the idea that the limits were set by the chance fixation of unfavourable alleles at the loci that were segregating in the base populations. Bearing in mind Robertson's (1960) derivation of the relationship between the half-life and the chance of fixation, the values observed for the half-life were sufficiently small to accommodate some margin of error in their estimation and still make the above statement valid. In other words, the selection as practised seems to have accomplished what it could reasonably be expected to accomplish, given these populations. A contribution to this end was undoubtedly the fact that the proportion of animals selected (about one-third) was close to the optimum, from the point of view of achieving the greatest possible advance. Robertson (1960) establishes that the maximum gain corresponds to a proportion selected of one-half; however, as the number of animals measured rises to 50 or so (as it did in the experiments discussed in this paper) the plot of limit against proportion selected becomes very flat topped, and the loss of potential gain by selecting only a third of the measured animals is but barely detectable. Fortuitously perhaps, the experiments discussed here seem to have featured high initial responses to selection without a sacrifice of ultimate gain, if we can safely conclude that unfavourable alleles have not been fixed. To combine these two objectives appropriately is a problem in practice, and one that has proved intractable to theoretical treatment.

The experiments reviewed in this paper seem to agree reasonably well with a model of selection limits based on the exhaustion of the additive genetic variance. It is emphasized however that this does not necessarily establish that model as the exclusive explanation of the phenomena. The genetic nature of the limits can be exposed to experimental investigation, as discussed in the next paper in the series.

SUMMARY

1. The results of some selection experiments for body weight in the mouse, onducted in the past in this laboratory, have been examined from the point of iew of the limits ultimately reached.

2. The limits that are apparently attained do not necessarily remain stable over rolonged periods of time; two large lines showed marked decreases despite con-nued selection for high body weight.

3. Selection for high body weight reached a limit in the region of 30 g. at 6 weeks f age; small mice reached their limit at around 12 g.

4. The time taken to reach the limit may vary from ten to thirty generations, even or this one trait.

5. The total response for unidirectional selection was between two and six times ne phenotypic standard deviation, or three to twelve times the additive genetic tandard deviation.

6. Consideration of the half-life of the selection responses excluded the likelihood f the chance fixation of alleles unfavourable to the direction of selection.

7. The loci contributing to the response could each have an effect amounting o anything from one-half to one phenotypic standard deviation in the base opulation.

8. This indicated that up to twenty loci had contributed to the response.

9. The intensity of selection practised was close to the optimum for obtaining ne maximum total response.

10. The rule of parsimony would indicate the exhaustion of the additive genetic ariance as an adequate explanation of the limits attained.

I should like to acknowledge the profit and pleasure of discussions with Drs D. S. Falconer, lan Robertson and W. G. Hill on various issues that arose during the preparation of this anuscript.

Dr Falconer kindly provided me with data to supplement his original publications. This cilitated greatly the examination of several points.

REFERENCES

EMPSTER, E. R. (1955). Genetic models in relation to animal breeding problems. *Biometrics*, **11**, 535–536.

ALCONER, D. S. (1953). Selection for large and small size in mice. *J. Genet.* **51**, 470–501.

ALCONER, D. S. (1955). Patterns of response in selection experiments with mice. *Cold Spring Harb. Symp. quant. Biol.* **20**, 178–196.

ALCONER, D. S. (1960*a*). *Introduction to Quantitative Genetics*. Edinburgh and London: Oliver & Boyd, Ltd.

ALCONER, D. S. (1960*b*). Selection of mice for growth on high and low planes of nutrition *Genet. Res.* **1**, 91–113.

ALCONER, D. S. & KING, J. W. B. (1953). A study of selection limits in the mouse. *J. Genet.* **51**, 561–581.

OODALE, H. D. (1938). A study of the inheritance of body weight in the albino mouse by selection. *J. Hered.* **29**, 101–112.

OODALE, H. D. (1941). Progress report on possibilities in progeny-test breeding. *Science, N.Y.* **94**, 442–443.

HILL, W. G. (1965). Studies on artificial selection. Ph.D. Thesis, University of Edinburgh.
HILL, W. G. & ROBERTSON, A. (1966). The effect of linkage on limits to artificial selectio *Genet. Res*, **8**, 269–294.
KIMURA, M. (1957). Some problems of stochastic processes in genetics. *Ann. math. Statist.* 2 882–901.
KING, J. W. B. (1950). Pygmy, a dwarfing gene in the house mouse. *J. Hered.* **41**, 249–252
MACARTHUR, J. W. (1944). Genetics of body size and related characters. I. Selecting small a large races of the laboratory mouse. *Am. Nat.* **78**, 142–157.
MACARTHUR, J. W. (1949). Selection for small and large body size in the house mous *Genetics*, **34**, 194–209.
NEWMAN, J. A. (1960). Reciprocal recurrent selection for body size in the mouse. Ph.D. Thes University of Edinburgh.
ROBERTSON, A. (1960). A theory of limits in artificial selection. *Proc. R. Soc.* B, **153**, 234–24

PAPER 6

The limits to artificial selection for body weight in the mouse. II. The genetic nature of the limits.

Genet. Res. Camb. 8, 361-375. 1966.

by

R.C. ROBERTS

Genet. Res., Camb. (1966), 8, *pp.* 361–375
With 7 text-figures
Printed in Great Britain

The limits to artificial selection for body weight in the mouse

II. THE GENETIC NATURE OF THE LIMITS

By R. C. ROBERTS

A.R.C. Unit of Animal Genetics,
Institute of Animal Genetics, Edinburgh, 9

(*Received* 10 *August* 1966)

1. INTRODUCTION

The first paper of this series (Roberts, 1966) examined the results of some earlier selection experiments for body weight in the mouse, conducted in this laboratory. The object was to determine the limits to selection that had been attained for this trait. Theoretical considerations of the experimental results led to the conclusion that the limits observed were compatible with a model based on the exhaustion of the additive genetic variance by fixation of loci contributing to the variation in weight. It was emphasized, however, that while no other explanation was necessary, other explanations were not specifically excluded by the analysis applied. This paper reports an experimental investigation of the genetic nature of the limits found in two of the selected lines included in the earlier study. The investigation establishes that in at least one of the two lines, total fixation was not an adequate explanation of the limit reached.

2. MATERIALS AND METHODS

It would obviously have been desirable to extend the experimental analysis of the limits to all seven lines described in the first paper. Unfortunately, the available cage space permitted only two of the lines to be studied in the way described here.

The two lines chosen for further study were the *CRL* and *CFS* lines, representing the large and small mice, respectively. These lines were first described by Falconer (1960), and their further progress under Dr Falconer's care was summarized in the first paper of this series. Both lines had been selected on growth between 3 and 6 weeks for thirty-one generations when I acquired them, and I am indebted to Dr Falconer for making these lines available to me.

The choice of these particular lines in preference to the others was governed by the following considerations. The *CRL* line was the largest of the four large lines that were available, and as it was to serve as a standard against which to assess methods of transcending the limit, it seemed a logical choice for a more detailed genetic study. The small *CFS* line did not meet this criterion so well, as it was the

largest of the three available small lines. However, the reproductive performance of the two smaller lines was by that time so poor that further work on them, which would require expansion of the stocks, presented serious practical difficulties. So the *CFS* line was preferred for continuation, as it had a better reproductive performance and also the added advantage of stemming from the same base population as *CRL*, the large line chosen. This would render any comparison between the large and small lines more meaningful.

Selection was continued in both of these lines, but from generation 32 onwards, the character selected was changed from growth between 3 and 6 weeks to 6-week weight itself. This was done for the sake of convenience, as Falconer (1955) had shown that the ranking of mice on the two measurements was virtually indistinguishable. But with the change in the selection procedure, the designations of the two lines were changed: *CRL* now became known in the laboratory as the *CL* line, while *CFS* became *CS*. This avoids confusion with Falconer's earlier (1960) study of these lines, while it also simplifies the designation of sublines drawn from the lines, as explained below. Frequent reference will be made to the *CL* and *CS* lines throughout the remainder of this series of papers. The selection was continued on a within family basis for a further twenty generations and more in each case; the sequential numbering of the generations was not broken.

Two offshoots were taken from each of the *CL* and *CS* lines. In one case, all selection was suspended, and the sublines became known as *CLR* and *CSR*, where the *R* stands for 'relaxed' selection. In the second pair of offshoots, the direction of the selection was reversed, and the sublines were called *CLB* and *CSB*, where the *B* stands for 'back' selection. In other words, *CLB* was the large line now selected for low body weight, while *CSB* was the small line selected for high body weight.

CLR was drawn at random from the 38th generation of *CL*. The remainder of the mice in that generation were selected as appropriate either to continue the *CL* line or to form the 1st generation of *CLB*, respectively. Similarly, *CSB* was derived from the 35th generation of the *CS* line while *CSR* was drawn at random from the 37th generation of *CS*. All of the lines were run on fifteen pair matings per generation.

3. RESULTS

(i) *Continued selection for body weight*

Though the character to which selection was applied was changed formally from post-weaning growth to weight at 6 weeks, there is no reason to suppose that the *CL* and *CS* lines had not attained the limit for 6-week weight by the time that I acquired them. This is amply confirmed in Fig. 1 which shows the progress of the two lines under selection from the time that they were formed. The dotted parts of the graphs summarize the weights up to generation 31, as discussed in the earlier paper. The solid lines represent the weights during the present study; these parts will be reproduced on a different scale in further figures, for the purposes of comparison with other studies, throughout the remainder of this series of publications.

Let us consider first the *CL* line. Under continued selection for 6-week weight, this line remained at much the same level as before (about 32 g.) for a further twelve generations. However, between generations 43 and 44, there was a marked increase in the mean body weight to 35 g., and the line has remained at this higher level for a further ten generations. Except for one sporadically high point at the 21st generation, the *CL* line is now running at a level that is clearly different from what it was before, after it had reached an apparent limit. The *CS* line, on the other

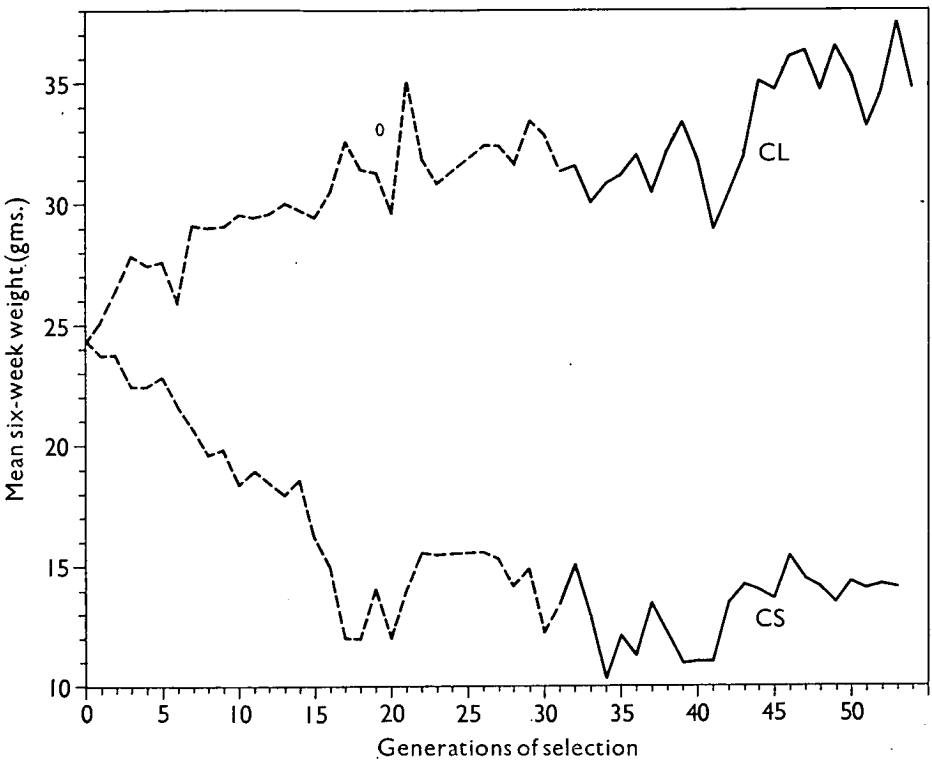


Fig. 1. Long-term responses to selection of *CL* and *CS* lines. Solid parts—weights during present study. Dotted parts—earlier responses, for comparison.

hand, does not arouse much suspicion of any major shift of a permanent nature in mean weight. Though its weights fluctuate over a range of about 5 g., there is no reason to revise the figure of 14 g. that was derived as the limit for this by generation 31 (see earlier paper).

It is important to decide whether the higher mean weights of the *CL* line from generation 44 onwards was of environmental or of genetic origin. Two factors might suggest an environmental cause. One is that the small (*CS*) line also showed an increase in weight at nearly the same time, i.e. at generation 42 of *CS*, which was contemporaneous with generation 43 of *CL*. The other reason is that the composition of the diet was modified at about that time, the modified diet being introduced

as *CL* mice of generation 43 and *CS* mice of generation 42 were approaching 6 weeks of age. There are, on the other hand, several reasons which cumulatively render it extremely unlikely that the change of diet was the cause of the increase in the *CL* line. Firstly, the modification of the diet was only slight, the main item being the replacement of miller's offal by ground wheat; the intention was to acquire a greater constancy of diet rather than to alter its nutritive value. Secondly, the increase in the *CS* line was not great compared to previous fluctuations in this line, and the increase occurred the generation before the large increase in *CL* was observed. Thirdly, no parallel increase in weight was shown by four other lines all derived from the *CL* line a few generations earlier (two of these lines are shown in Fig. 4). A dietary effect on weight would therefore have to be highly specific to the *CL* line, and from a recent review of the literature (Roberts, 1965) it would seem very improbable that genotype–environment interactions of this magnitude should appear in strains separated by only a few generations. For these reasons, it seems much more likely that the increase in weight of the *CL* line was of genetic origin, and this interpretation will be adopted in the evaluation of the data discussed in the remainder of this series of papers. It should be noted that the change occurred well after various offshoots of the *CL* line were propagated for other studies.

The genetic nature of the change in weight is open to several interpretations. The fact that it occurred, over three generations, after a depression of the mean weight to its lowest level for over thirty generations accords well with the model of selective peaks, separated by 'saddles', often expounded by Wright, and recently (1965) reviewed by him. This model is a fairly complex one whose main feature is genic interactions. At a simpler level, we could postulate a new mutation favourable to the direction of selection. Perhaps the most likely model would invoke a rare recombinational event as suggested by Thoday & Boam (1961), and confirmed by Thoday, Gibson & Spickett (1964), to explain similar shifts in the mean bristle counts of selected lines of *Drosophila*. With an organism like the mouse, there is at present little hope of being able to distinguish between these various models experimentally. The main point is that we do not lack plausible genetic explanations of the increase in weight of the *CL* line that are still consistent with a temporary limit to selection resulting from the exhaustion of the original additive genetic variance.

To summarize, two conclusions emerge from this section. Firstly, there is no evidence that a formal change in the character selected in any way affected the conclusions drawn earlier (Roberts, 1966) with respect to the limits of artificial selection for body weight in these two lines. Secondly, we have now a third instance (out of four possible cases) of a large line proving to be unstable with respect to body weight after it had apparently reached a limit. But whereas the two cases reported in the earlier study both showed a shift contrary to the direction of selection, and remain completely unexplained, the one reported here was in the direction of selection and is much more open to an acceptable genetic interpretation. One may ask, however, whether the shifts in all three lines may not have been different facets of a common phenomenon. In terms of Wright's model, is it possible that the two

cases which showed a decline in weight accidentally drifted to a lower peak? An opportunity for this to happen would stem from a lowering of selection pressures, which might result from the reduced fertility that frequently characterizes highly selected lines.

(ii) *Test for additive genetic variance in lines at the limit*

If all the loci affecting body weight in a line have been fixed by selection, then of course there will be no genetic variance of any description left in that line. This, however, is not very easy to test without special experimental programmes. But existing data can be utilized to see whether there is any heritable (or additive) variance available. If additive genetic variance is present, then this ought to be reflected in correlations between relatives which would lead to positive estimates of the heritability.

Maternal effects on body weight in the mouse are well known, and since these grossly affect estimates of genetic parameters (Falconer, 1964), relationships such as full sibs that involve a common dam are not of much use for present purposes. Regressions on sire, however, do not incur these complications, and although the number of sires used in any one generation was at the most fifteen, different generations can be pooled to obtain reasonably accurate estimates of the regression of offspring on sire. Data from generations 32 onwards have been employed to obtain estimates of the heritability of body weight in the *CL* and *CS* lines after they had reached their limits, with the following results:

$$CL \text{ line: Heritability} = 0{\cdot}194 \pm 0{\cdot}120$$
$$CS \text{ line: Heritability} = 0{\cdot}180 \pm 0{\cdot}092$$

The estimates of the heritability in the two lines are very similar in absolute terms. The estimate for the large line is not significantly different from zero, but that for the small line is on the borderline of formal statistical significance at the 5% level. This suggests that, in the small line at least, a substantial proportion of the variance in weight may be additive genetic. If this is so, it follows that the limit in the *CS* line is not fixed by the exhaustion of the genetic variance. This point will be amplified in section (iv) below.

(iii) *Relaxed and reversed selection in the large line*

The first effect of the cessation of selection for large size was a practical one that became immediately apparent, namely, that on account of their increased fertility, both the *CLR* and *CLB* lines became much easier to maintain than the parent *CL* line, in which selection for high 6-week weight was continued. Two aspects of this increase in fertility are summarized graphically in Figs. 2*a* and 3. Figure 2*a* gives the frequency distribution of the number of fertile matings, out of the fifteen that were set up in each generation in each line over the period of study. With two trivial exceptions, the *CLR* line regularly equalled or excelled the *CL* line over the sixteen generations that have elapsed since its formation. Its fertility on this

measure has been consistently high. The *CLB* line, on average, fell in between the other two, but it also never caused any concern on account of infertility. Figure 3 shows the number of live young at birth in the first litters on the three lines. Over the early generations, especially, the superiority of both the *CLR* and *CLB* lines over *CL* is unquestionable; in later generations, this superiority in litter size faded away.

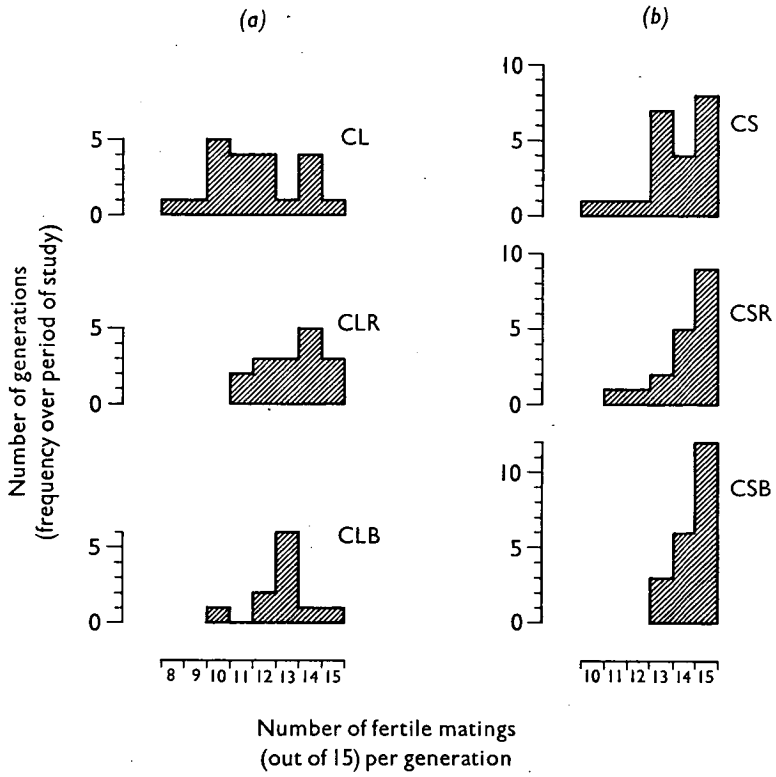


Fig. 2. Frequency distributions of the number of fertile matings (out of fifteen) in *CL* and *CS* lines from generation 32 onwards, and their derivatives.

While this increased fertility of the relaxed- and back-selection lines was a welcome feature on managemental grounds, it posed a problem with respect to the genetic interpretation of any changes in body weight. There is a well-known maternal effect on body weight in the mouse, as mice gestated and reared in large litters have their weights depressed as a consequence. In order to compare mice of different lines, it is therefore desirable to adjust the weights to a common litter size. By pooling data within generations, the expected negative regression of 6-week weight on number born was found in all lines, the depression of mean weight being in the region of 0·6 g. for each extra mouse born. This is considerably higher than the value of 0·34 found by Falconer (1964) in an unselected strain of the same origin. However, when the regression was calculated from the generation means, there was not the slightest evidence that the same relationship held. Using the generation

means to regress body weight on litter size is not very accurate, as the degrees of freedom are limited, but cumulatively, the lines reported here and some others showed no hint of any consistency in the sign of the regression coefficients which, of course, individually never even approached statistical significance. It was therefore decided that the weights should *not* be corrected for litter size. This decision was helped by the fact that the within-generation regressions, when these were applied to generation means as a trial, generated such small corrections that no conclusion could possibly be affected.

The mean 6-week weights, plotted against generation number, are shown in Fig. 4. Disregarding the change in weight of the *CL* line, discussed earlier, neither

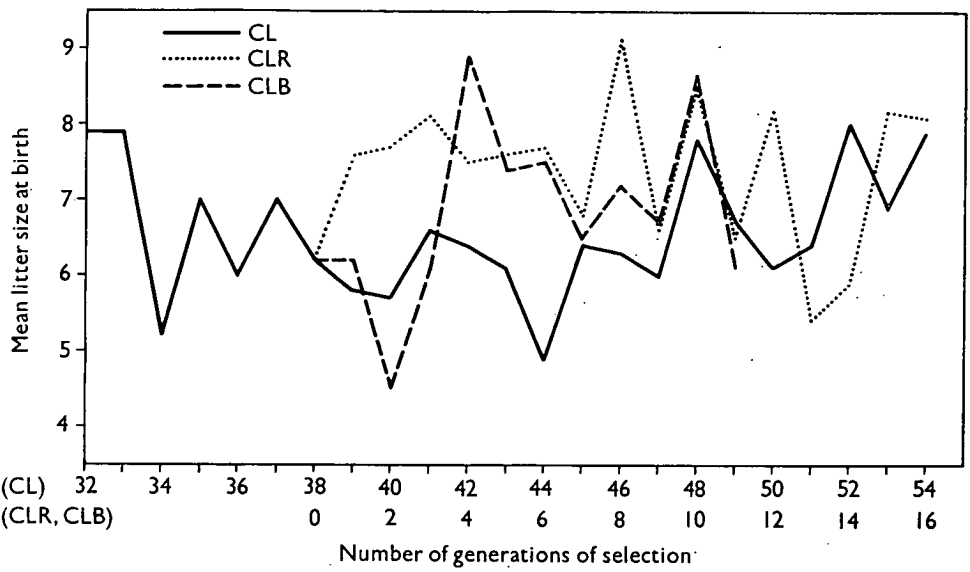


Fig. 3. Effects of relaxed and reversed selection in large line on number of live young at birth in first litters.

relaxed nor even reversed selection had any effect on body weight, and both sublines continued at precisely the old level of the *CL* line. The *CLB* line, the one selected for a reduction of body weight, accumulated a total selection differential of 15 g. over its eleven generations. When, in its 11th generation, it showed its highest-ever weight, and it had quite obviously failed to show any response to downward selection, it was discontinued. The *CLR* line was continued for other reasons; over its sixteen generations, it has accumulated a negligible positive selection differential of less than 2 g., pointing to the successful randomization of animals chosen as parents.

The heritabilities in the *CLR* and *CLB* lines were calculated from the regression of offspring on sires, pooled within generations, with the following results:

$$CLR \text{ line: Heritability} = -0{\cdot}016 \pm 0{\cdot}098$$
$$CLB \text{ line: Heritability} = +0{\cdot}084 \pm 0{\cdot}158$$

These estimates obviously do not differ from zero, and confirm the conclusion from the back selection that no additive genetic variance remained in the *CL* line.

The conclusion from this section is clear: there was no additive genetic variance of body weight remaining in the *CL* line at the selection limit. The absence of additive variance, however, is not quite synonymous with the fixation of all loci affecting body weight. Overdominance could, in principle, lead to a situation where some loci segregated without showing any additive variance, if heterozygotes were selected on the basis of weight alone. Gene frequencies would then equilibrate at some intermediate level, and only after an accidental deviation from equilibrium

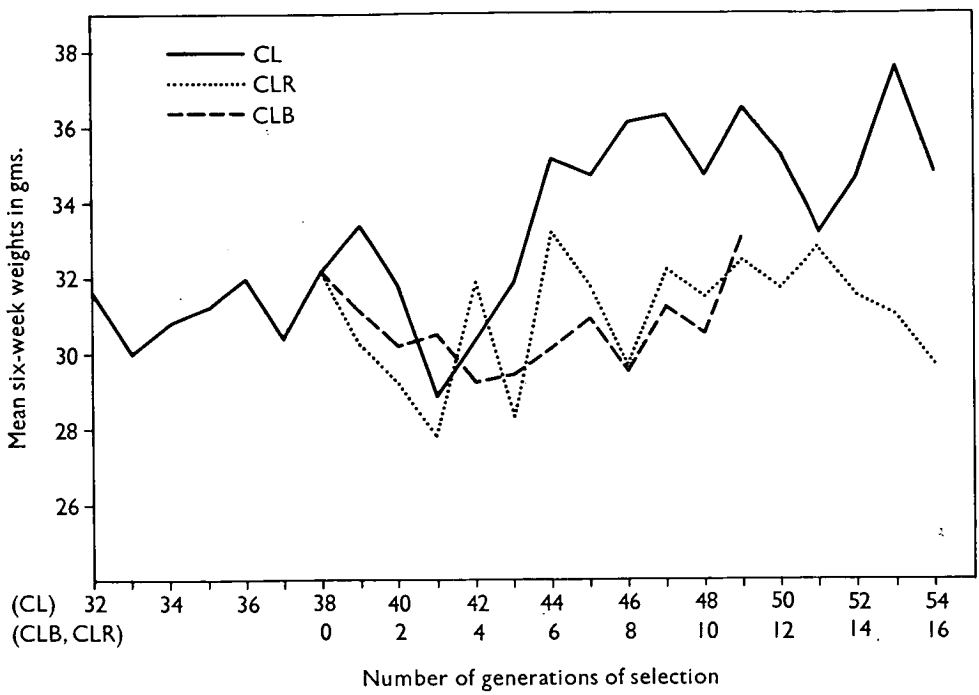


Fig. 4. Effects of relaxed and reversed selection in large line on mean 6-week weight.

could a response be obtained to reversed selection. Overdominance of this kind could be simulated by close linkage between pairs of genes; recombination between the members of such a pair was invoked earlier as a likely explanation of the increase in weight of the *CL* line. Overdominance of another kind was specifically excluded by the experimental results. This situation demands that natural selection should oppose the artificial selection; one homozygote would be rejected because of its effect on weight, while the other homozygote would tend to be either infertile or inviable, in which case the heterozygote might be fitter than either homozygote, under the conditions of the experiment. But under these conditions, while there would be no additive variance of overall fitness, there would be additive genetic variance of body weight. The fact that weight did not change when selection was

relaxed or reversed proves that if there were any overdominant genes, or linked pairs, still segregating, they caused a negligible amount of variance.

From the fact that all relevant loci were fixed, we can deduce something about the cause of increased fertility which, as mentioned earlier, followed the cessation of selection for large size. Since no genetic variance in body weight remained, this must mean that there was a negative environmental correlation between large size and fertility. The productivity of large mice will be examined in more detail, and in a more appropriate context, in a future paper.

(iv) *Relaxed and reversed selection in the small line*

When selection for low body weight was stopped, the effects on the fertility of mice of the *CS* line were again beneficial. The number of sterile matings fell, as shown in Fig. 2*b*, the reduction being more noticeable in the back selected (*CSB*)

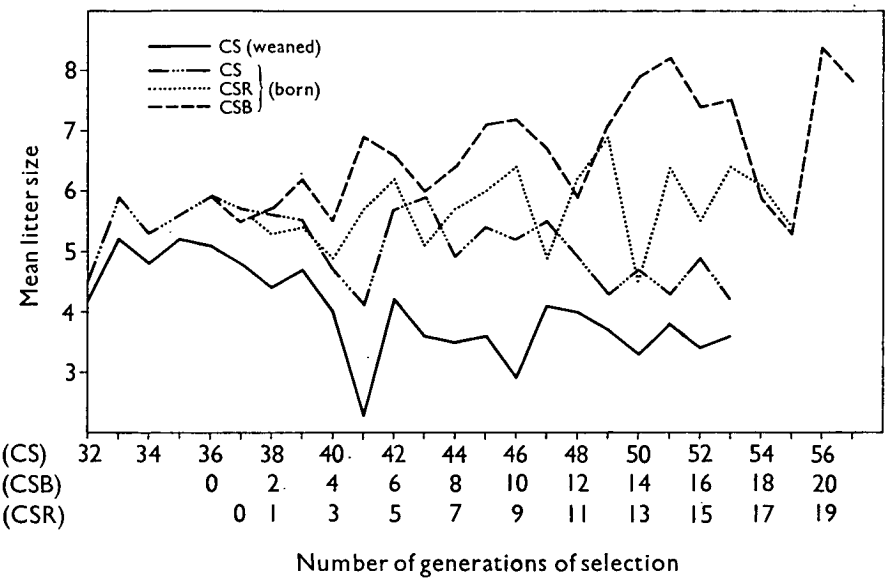


Fig. 5. Effects of relaxed and reversed selection in small line on number of live young at birth in first litters. Number weaned in *CS* line also shown.

line than in the relaxed (*CSR*) line. This is slightly different from the large lines, where the ranking of the relaxed and back-selected lines was reversed. As a general point, we may note also that sterility is commoner in large mice than in the corresponding small ones. The other criterion of fertility that was examined, namely the mean litter size at birth, also showed an improvement in both lines, and again the increase was more conspicuous in the *CSB* line than in the *CSR* line, as shown in Fig. 5. But for reasons given previously, these litter size differences were not employed to attach adjustments to body weights.

The effects of the relaxed and reversed selections on body weight are summarized in Fig. 6. Some violent changes in weight occur in all three lines from time to time; as usual in similar situations, some of the weight changes in different lines are

synchronous, while others are not. However, despite the fluctuations, a fairly clear picture emerges. The relaxed line (*CSR*) gives a hint that it may have increased in weight slightly but there is no indication that the difference between it and its parental line (*CS*) has increased at all with time. The average difference between the two lines has been of the order of 1 g. or so for the last ten generations. The line in which the direction of the selection was reversed (*CSB*) shows quite a clear-cut result. Progress was made regularly over about thirteen generations of reversed selection, and the mice, weighing by now about 17 to 18 g., are no longer particularly small. But for the last eight generations, no further progress has been made, suggesting that probably a limit has now been reached for the reversed selection.

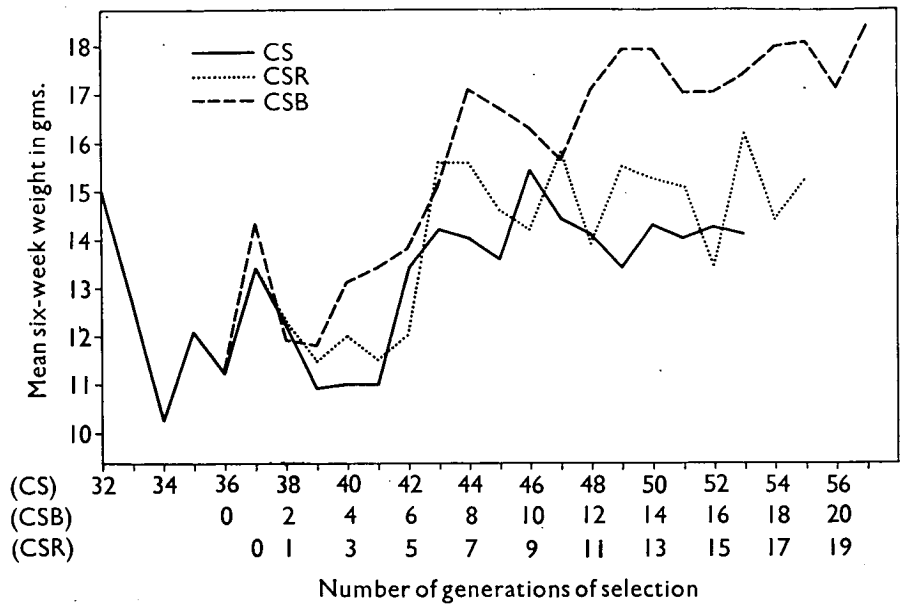


Fig. 6. Effects of relaxed and reversed selection in small line on mean 6-week weight.

Whether a limit has been reached or not, there is no doubt about the reality of the response to reversed selection. This accords with the heritability estimate given earlier, and proves that, unlike the corresponding large line the additive genetic variance in the small (*CS*) line had not been exhausted by thirty-six generations of selection for small size. For the previous twenty generations or so, the *CS* line had reached an apparent limit to downward selection, which indicates that some process had prevented at least some of the alleles affecting body weight from going to fixation.

The heritability estimated from the offspring-sire regression in the *CS* line was given earlier as $0{\cdot}180 \pm 0{\cdot}092$. Estimates derived in similar fashion from the *CSR* and *CSB* lines were the following:

$$\textit{CSR} \text{ line: Heritability} = 0{\cdot}080 \pm 0{\cdot}078$$

$$\textit{CSB} \text{ line: Heritability} = 0{\cdot}156 \pm 0{\cdot}088$$

Though the estimate from *CSR*, on its own, is insignificant, the estimate from *CSB* agrees well with the one originally obtained from *CS*, while all three are consistent within the limits of their sampling errors.

It is well known that the heritabilities realized in practice may differ markedly from estimates derived in this manner. To test this, the heritability realized by selection in the *CSB* line was computed from the data summarized in Fig. 7, which shows a plot of generation means against the cumulated selection differential. The

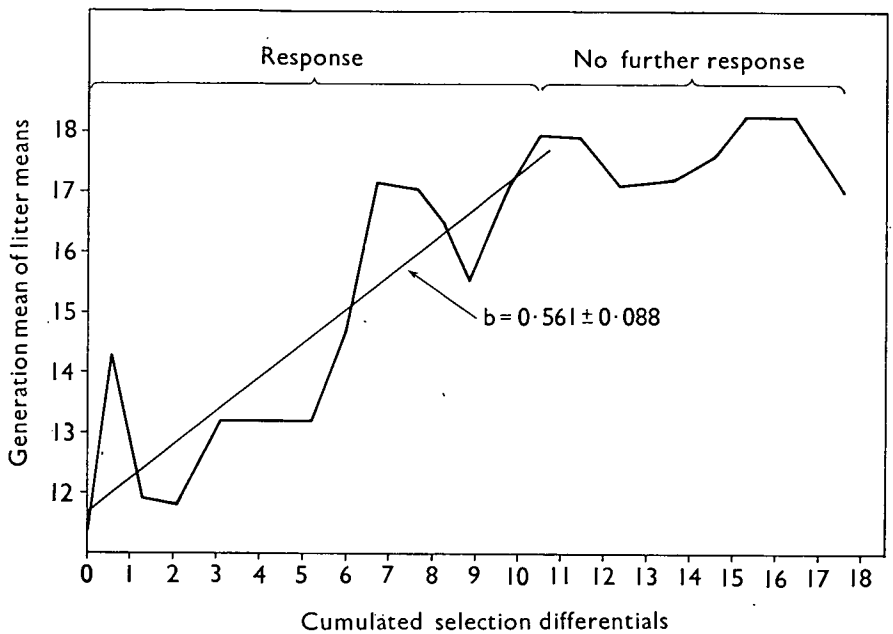


Fig. 7. Mean 6-week weight in *CSB* line plotted against cumulated selected differential. The straight line is the least squares regression line over the period of the response.

generation means have been calculated in a way slightly different from the ones given previously; this time, they are the means of litter means. This is a more appropriate measure for present purposes, as the selection differentials were calculated (the method of selection being within—families) as the mean deviation of selected animals from the means of the litters in which they were measured. Selection differentials were calculated separately for the two sexes and averaged, sterile animals being disregarded. The graph has been divided somewhat arbitrarily into two parts, one corresponding to the period of initial response (thirteen generations), and the other to a second period when there was no obvious response and which we suppose represents the limit to the reversed selection. Only the first part of the graph was utilized to calculate the realized heritability, which turned out to have the surprisingly high value of $0{\cdot}561 \pm 0{\cdot}088$. This value is probably inflated somewhat. *CSB* was taken off the 36th generation of *CS*, and as can be appreciated best from Fig. 1, this represents a low value for the *CS* line. Some of the apparent response,

in absolute terms, of the *CSB* line may therefore have been due to environmental increases that occurred later. These difficulties could have been circumvented, in principle, by recording the response of the *CSB* line as a deviation from *CS*. Unfortunately, there was not sufficient contemporaneity between generations in the two lines to render this a satisfactory procedure. But if, instead, we take the final superiority of the *CSB* line over *CS* to be 3·5 g. (from Fig. 6), and that this had been achieved by the time the cumulated selection differential was 10·5 g. (from Fig. 7), it now gives us a realized heritability of 33%. Though well within the upper confidence limit of the original estimate obtained from *CS* line, it is still high when we consider that the additive genetic variance in the *CS* line, though obviously not exhausted, must surely have been severely depleted.

To what, then, must we ascribe the preservation of so much additive genetic variance in the *CS* line? An obvious factor to test is natural selection, but it becomes very difficult indeed to test this factor adequately. There are some indications that natural selection may oppose the artificial selection; the slight increase in weight noted when selection was relaxed could be interpreted in this way. It was also noted that fertility and litter size increased on relaxation, but there is no evidence that the sterility in the *CS* line has had any adverse effect on the selection process. This can be determined from a comparison of the expected and realized selection differentials. The expected selection differential is the mean superiority of selected individuals. The realized selection differential, for a within-family method of selection, is the superiority of animals that proved themselves fertile, and therefore had a litter measured in the next generation. The expected selection differential cumulated over the last twenty generations of the *CS* line was only 0·28 g. greater than the realized, or 0·014 g. per generation. This proves conclusively that natural selection did not operate through any differential fertility between the smallest mice and those not quite so small. This, however, takes no account of any natural selection that may have operated on viability between conception and the time when the animals were measured at 6 weeks of age. It is not possible, from the data available, to estimate the selection differential that may have been lost on account of mortality. It is, however, common laboratory experience that the losses are heaviest among the smallest mice within any one stock; indeed, the phenomenon is by no means confined to the mouse alone. Some indication of the extent of the mortality is shown in Fig. 5, which shows both the number born alive and the number weaned in the *CS* line. The loss between birth and weaning is frequently 20%, and sometimes 40%. This, of course, represents only a part of the total mortality. Much earlier in its history (generations 16 to 19), the reproductive performance of the *CS* line was examined by Fowler & Edwards (1960), who reported losses of 28% between ovulation and the recording of live births. The performance of the line would certainly not be expected to have improved since that time. It is therefore probable that fully half of the number conceived die before weaning time. Further losses occur before selection at 6 weeks of age. Over its last twenty generations, the *CS* line showed a mortality rate of 11% between 3 and 6 weeks of age, though the mean weaning weight of the animals that died was only about a tenth of a gramme less than the

weaning weight of the survivors. This apparently slight reduction, however, may mean a much greater reduction of selection differential in 6-week weight; there are complex relationships between the variance components of weights at successive ages, as discussed by Monteiro & Falconer (1966). As the total selection differential obtained was only about 0·6 g. per generation, any tendency for the smaller mice to die in the post-weaning phase may contribute a substantial proportionate effect on the selection differential.

If the failure to respond to artificial selection is to be attributed to the opposing effect of natural selection, acting through differential viability, then the mice selected among the eventual survivors must have a mean weight equal to what the mean weight would have been had all zygotes survived. In other words, the positive deviation of those rejected by artificial selection must have been counterbalanced by the negative deviation of those that failed to survive to 6 weeks. In view of the heavy mortality, and the fact that about a third to a half of the survivors were selected, this seems to be a reasonable postulate.

Though the evidence is only indirect, it seems justifiable to conclude that the limit to selection for small size in the *CS* line is due to natural selection opposing the artificial selection, and that further the natural selection operates through its effect on viability and not on the fertility of the survivors. The one slight difficulty is that, if the hypothesis of reduced viability is correct, the weights should increase so little when the artificial selection was relaxed. It appears as if the natural selection may not exist at all until body weight is reduced to some particular level. Some support for this idea may be derived from Fig. 1. The initial response to selection in the *CS* line was rapid and, if not linear, accelerating. But at generation 17 or so, it came to an abrupt halt, and failed to show any further response over the next thirty-five generations. This suggests strongly that the barriers to further progress were encountered at a particular weight, but that none of their effects were felt until that weight was reached. This is strongly reminiscent of F. W. Robertson's (1963) finding in *Drosophila*, that there is a critical larval weight below which pupation fails to occur. It is not wildly speculative to suppose that some analogous phenomenon may exist during the development and growth of the mouse, and that in the *CS* line, this critical low weight may have been reached.

4. DISCUSSION

It was seen in the preceding sections that the limit to artificial selection had been reached for very different reasons in the large and small lines. In the large line the additive genetic variance had been effectively exhausted. In the small line, however, a substantial proportion of the remaining variance was additive genetic, and a response to reversed selection was readily obtained.

It was explained earlier that only two of the seven selected lines available for study were subjected to further experimental investigation of the nature of the limits. However, Falconer (1955) reports some short-term studies of a similar kind

on two of the other five lines. Reversed selection was carried out from the small (*NS*) line on two separate occasions. The first (from generation 12) was at a time when the line was still responding, but by the second time (from generation 20) the line was approaching its ultimate limit. Over four generations, the response to the reversed selection was unmistakable. The other study described by Falconer was the relaxation of selection from the 24th generation of the large (*NF*) line, after the line had reached its limit. Over six generations, there was no indication that the relaxation of selection resulted in any separation from the line under continued selection.

Though the evidence just quoted is fragmentary, it does encourage some thought of the possible generality of the phenomena described in this paper, with respect to selection for body weight in the mouse, namely that selection for large size may lead to the exhaustion of the additive genetic variance whereas selection for small size may reach a limit despite the detectable presence of additive variance. If this is so, then the genetic nature of the limits were reversed from the ones that appear to obtain in *Drosophila*; in this organism, it is selection for *small* size that seems to lead to fixation. Reeve & F. W. Robertson (1953) described a strain, selected for fifty generations for long wings, in which the additive genetic variance was much greater than in the base population and from which relaxed and reversed selection yielded ready responses. F. W. Robertson (1955) reported a parallel but extended study, using thorax length as his criterion of size. After twenty generations of selection, the small flies failed to yield any response to further selection in either direction. The large flies, on the other hand, reached the limit to further selection after twelve to fifteen generations but quickly returned to the level of the base population on the reversal of selection. Detailed analyses in both of these *Drosophila* studies indicated to the authors that genetic mechanisms of some complexity operated to preserve heterozygosity in the lines selected for large size.

Another *Drosophila* study on the long-term effects of selection, this time for a bristle score, was reported by Clayton & A. Robertson (1957). Despite the highly additive genetic basis of the character selected, a limit to the response in either direction was still compatible with a considerable amount of residual genetic variance. In their high lines, the variability was attributable to the continued selection for lethal heterozygotes. In the low lines, the situation appeared to be particularly complex, lethal genes, infertility of extreme females and inversion heterozygotes all being invoked to explain some of the residual genetic variance.

The results so far available on selection limits suggest that models based on the exhaustion of the additive variance may not be sufficiently comprehensive to describe fully many of the situations derived in practice. They therefore underscore the need for more detailed investigations of specific cases, if we are to gain a deeper appreciation of the genetic nature of the limits to artificial selection. This objective may be less remote if organisms showing some diversity of biological organization could be included in such studies, which furthermore ought to include characters cast in different evolutionary moulds, if any generalities are to emerge.

SUMMARY

1. The effects of long-continued selection for body weight in two lines of mice, one large and one small, are described.

2. The large line showed a sharp increase in weight after remaining at an apparent limit for twenty generations. A rare combinational event is suggested as the most likely explanation.

3. Reversed and relaxed selection from the large line at the limit failed to yield any response. This indicates that effectively, the additive genetic variance in this line had been exhausted.

4. In contrast, the small line at the limit regressed slightly towards the base population when selection was relaxed. Reversed selection yielded a ready response until a new limit was apparently reached. Loci affecting body weight in this line had therefore not been fixed by selection.

5. Natural selection, operating on viability between conception and the time when the selection was made, appears to explain best the lack of fixation in the small line.

6. Attention is drawn to the necessity of more experimental work to elucidate the genetic nature of the limits to artificial selection.

I am much indebted to Dr B. Woolf for the emergency provision of a computer programme to cope with some of the statistical analyses on which this paper is based.

Dr D. S. Falconer kindly suggested many improvements in the presentation of the material.

REFERENCES

Clayton, G. A. & Robertson, A. (1957). An experimental check on quantitative genetical theory. II. The long-term effects of selection. *J. Genet.* **55**, 152–170.

Falconer, D. S. (1955). Patterns of response in selection experiments with mice. *Cold Spring Harb. Symp. quant. Biol.* **20**, 178–196.

Falconer, D. S. (1960). Selection of mice for growth on high and low planes of nutrition. *Genet. Res.* **1**, 91–113.

Falconer, D. S. (1964). Maternal effects and selection response. In *Genetics today. Proc. XIth int. Congr. Genet. (The Hague)* 1963, Vol. III, 763–774.

Fowler, R. E, & Edwards, R. G. (1960). The fertility of mice selected for large or small body size. *Genet. Res.* **1**, 393–407.

Monteiro, L. S. & Falconer, D. S. (1966). Compensatory growth and sexual maturity in mice. *Anim. Prod.* **8**, 179–192.

Reeve, E. C. R. & Robertson, F. W. (1953). Studies in quantitative inheritance. II. Analysis of a strain of *Drosophila melanogaster* selected for long wings. *J. Genet.* **51**, 276–316.

Roberts, R. C. (1965). Some contributions of the laboratory mouse to animal breeding research. Part I. *Anim. Breed. Abstr.* **33**, 339–353.

Roberts, R. C. (1966). The limits to artificial selection for body weight in the mouse. I. The limits attained in earlier experiments. *Genet. Res.* **8**, 347–360.

Robertson, F. W. (1955). Selection response and properties of genetic variation. *Cold Spring Harb. Symp. quant. Biol.* **20**, 166–177.

Robertson, F. W. (1963). The ecological genetics of growth in *Drosophila*. 6. The genetic correlation between the duration of the larval period and body size in relation to larval diet. *Genet. Res.* **4**, 74–92.

Thoday, J. M. & Boam, T. B. (1961). Regular responses to selection. I. Description of responses. *Genet. Res.* **2**, 161–176.

Thoday, J. M., Gibson, J. B. & Spickett, S. G. (1964). Regular responses to selection. II. Recombination and accelerated response. *Genet. Res.* **5**, 1–19.

Wright, S. (1965). Factor interaction and linkage in evolution. *Proc. R. Soc.* B, **162**, 80–104.

PAPER 7

The limits to artificial selection for body weight in the mouse. III. Selection from crosses between previously selected lines.

Genet. Res. Camb. 9, 73-85. 1967

by

R.C. ROBERTS

Genet. Res., Camb. (1967), **9**, *pp.* 73–85
With 4 text-figures
Printed in Great Britain

The limits to artificial selection for body weight in the mouse

III. SELECTION FROM CROSSES BETWEEN PREVIOUSLY SELECTED LINES

By R. C. ROBERTS

A.R.C. Unit of Animal Genetics,
Institute of Animal Genetics, Edinburgh, 9

(*Received* 18 *August* 1966)

1. INTRODUCTION

Previous papers in this series (Roberts, 1966*a*, *b*) examined the limits to artificial selection for body weight in the mouse, and the genetic nature of those limits. It was found that, in a line selected for large size, the additive genetic variance had, effectively, been exhausted through the fixation of alleles contributing to large size. A line selected for small size, on the other hand, displayed a surprising amount of residual genetic variance at the limit, and it responded readily to reversed selection. However, this reversed response, as it tailed off, fell well short of the initial level of the base population, which established that the small line also had undergone a considerable amount of fixation, as indeed would be expected. Some evidence was adduced that these results may be representative of five other selected lines, three large and two small, that had been developed in this laboratory. This leaves open the question whether different lines selected in the same direction are fixed for the same alleles at the various loci affecting body weight. If they are not, then crosses between such lines ought to contain some genetic variance, and a response to further selection from the crosses may be expected. The limit to this second cycle of selection will depend on the extent of genetic differentiation between the lines at their original limits.

A precedent for this approach, with encouraging results, is reported by Falconer & King (1953). They obtained samples of two strains of mice selected for high 60-day weight, one by Goodale (1938, 1941) and one by MacArthur (1944, 1949). By the time the samples of these strains were procured, both had apparently reached a limit to selection, corresponding to a 6-week weight of about 29 g. in each case. Falconer & King noted that whereas Goodale's strain was large-bodied and not very fat, MacArthur's strain was smaller in linear dimensions but was very fat. From this observation, Falconer & King argued that a cross between them should provide new genetic variance upon which continued selection could act. This expectation was realized in practice, and over the nine generations of further selection which they reported, the mean weight rose by almost 3 g. to 32 g.

The work reported in this paper is an extension of Falconer & King's approach. The intention was to examine in more detail the potentiality of crossing selected lines to provide material for further selection, and to determine by how much the original limit to selection might be transcended.

2. MATERIALS AND METHODS

The experimental work described here stems from two base populations that were constructed from lines of mice that had been selected to the limit either for high or for low 6-week weight. The first population derived from four lines selected for high 6-week weight, and the second from three lines selected for low 6-week weight. A description of these seven original lines, and a report of the limits to selection which they had reached, is given in the first paper of this series (Roberts, 1966*a*).

The two base populations for the present studies were constructed as follows. As the scheme differed somewhat for the two, they are described separately.

Combining the four large lines presented no problem. They were first completely intermated according to a 4 by 4 diallel scheme, the 'pure lines' being included for comparison with the crosses. With the one exception noted below, the 'pure lines' were then discarded and a sample of each cross was mated to its complement, *i.e.* to a cross between the other two lines. Reciprocal crosses were included in all possible combinations. Each individual progeny of this generation thus had each of the four original large lines represented in its ancestry in equal proportions. This means that the gene frequencies at segregating loci had values of 0·25, 0·50 or 0·75. From the 120 matings that had been set up, fifteen fertile ones were chosen at random to provide a litter for continuing the stock, the random choice being disturbed only to ensure that different maternal combinations (and by reciprocity, the paternal ones) were represented as equally as possible. The fifteen litters so selected were designated the zero generation of the *LX* stock (*L* for 'large', and *X* for 'crosses').

Combining the three small lines was slightly more cumbersome; it is a consequence of diploidy that it is easier to combine four strains equally than three. The first step was exactly as before, and two-line crosses were extracted from a 3 by 3 diallel. Again, one 'pure line' was continued, while the other two were discarded. A random sample of each cross was then mated to a cross involving the third line, in all possible combinations. This, however, meant that mated animals shared one parental line in common. In the next generation, matings were between three-line cross animals, with the restriction that the common parental line should differ in the two mates. The progeny of this generation therefore had the three original lines represented in their ancestry in the proportions of 3:3:2, as far as an individual progeny was concerned. But as the crossing had been done comprehensively and schematically, in the population as a whole the three original lines were represented equally. Gene frequencies at segregating loci were thus either 0·33 or 0·67. From the seventy-two matings that had been set up, fifteen were chosen to provide the litters that

were designated the zero generation of the second base population constructed, *SX* (*S* for 'small', and *X* for 'crosses'). The choice of litters was at random from within subgroups, care being taken to maintain the equal representation of the original three small lines.

Having thus constructed the two base populations, they were thenceforth treated similarly. From the zero generations, the *LX* line was selected for high 6-week weight whereas the *SX* line was selected for low 6-week weight. Each line was maintained on fifteen pair matings, and the within-family method of selection was practised in both cases.

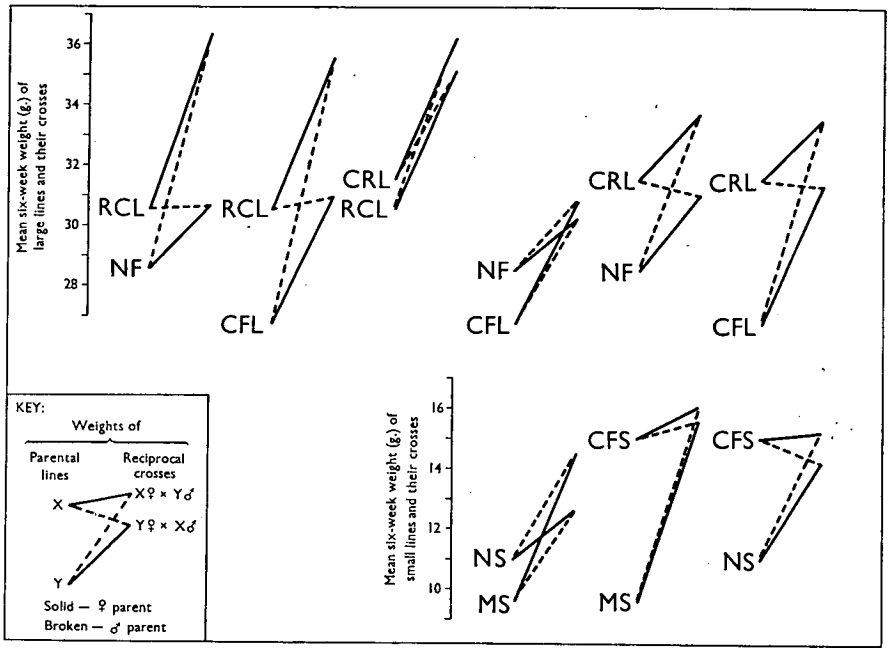


Fig. 1. Heterosis shown in crosses between selected lines of mice. Strain designations as given by Roberts (1966*a*).

For purposes of comparison with lines selected from crossbred material, one large line (*CL*) and one small line (*CS*) were maintained. These are reproduced in some of the figures with little further comment; they have been described fully in an earlier paper (Roberts, 1966*b*).

3. RESULTS

(i) *Differentiation between selected lines*

The detailed results of the line-crossing undertaken to form the base populations are not of great relevance in the present context. The most pertinent feature concerns the first stage of the crossing, and the results are summarized in Fig. 1. This shows the mean weights of the two-line crosses, the reciprocals being shown separately, compared with their 'pure-line' contemporaries from the two diallels. The

designations of the lines are those given by Roberts (1966*a*). For the large lines and their crosses, each point represents the mean of (usually) some twenty to fifty individuals, and have standard errors of somewhere between one-half and three-quarters of a gramme, depending on the number. The means for the small lines and crosses, by virtue of the lower variance of small mice, have standard errors about half as great. The important point for the present is that *all* crosses displayed considerable heterosis in body weight, at least one (and usually both) of the reciprocals exceeding the better parental line. The fact that all crosses did this means that all of the lines crossed were genetically differentiated to some degree with respect to body weight, for heterosis can result only from dominance or epistatic relationships between differing alleles. The increase in weight on crossing, even among the small lines, confirms the well-known fact that directional dominance favours a higher body weight in the mouse. In other words, genes for low body weight tend to be recessive.

The higher mean weight obtained when the lines were first crossed was not increased any more by further crossing, as can be seen from the summary of mean weights at different stages of the crossing shown in Table 1. However, some increases may have been obscured because the fertility of the two-line crosses (also shown in Table 1) was much higher than that of the parental lines, especially in the case of

Table 1 *Mean body weights and litter sizes of crosses between selected strains*

Stage of crossing (see text)	*LX* population (from large lines)		*SX* population (from small lines)	
	Litter size	6-week weight	Litter size	6-week weight
2-line cross	7·03	32·75	4·36	14·76
3 or 4-line cross	10·59	32·04	5·31	15·14
Further cross	—	—	5·56	14·66

the large mice. Although I have argued earlier in this series of papers (Roberts, 1966*b*) against the adjustment of generation mean weights for litter size differences, it is possible that an increase of 50% in fertility (as in the large mice) should not be ignored; it may have depressed the mean weight by, perhaps, 2 g.

During the formation of the base populations, there was therefore *a priori* evidence that new genetic variance would be available, because the selected lines that were employed for the crossing were differentiated genetically at loci contributing to variance in body weight. Furthermore, it appeared subjectively that this differentiation was widespread and pronounced. Even the two closely related large lines, *CRL* and *CFL*, drawn initially from the same source, showed the usual amount of heterosis on crossing. However, *CRL* was originally selected on a restricted diet, while *CFL* was selected on a full diet (Falconer, 1960), and as shown by Falconer, the genetic correlation between growth on the two planes, taking the average of four estimates, is only about 0·5.

The crossbred populations were formed to provide bases for further selection for large or small size, as appropriate. The results from these two operations are given separately.

(ii) *Further selection for large size*

From the zero generation, the *LX* line was selected for a further eighteen generations for high 6-week weight, after which it became extinct through infertility. The cause of the infertility appeared to be excessive fatness in females, few of whom ever gave birth to a second litter in the later stages of the experiment, and many of whom failed to produce even one litter. Males, on the other hand, when mated to females of more normal body size, were fertile for at least a few months. The trouble in the *LX* line arose when mating had to be delayed until sufficient animals reached 6 weeks of age, by which time the older females were 8 to 10 weeks old and were already too fat to breed. A later derivative of *LX*, which is not described further in this paper, was mated at 5 weeks of age, which did not permit an excessive accumulation of fat before mating. The early mating overcame the fertility problems in the line completely.

Before it became extinct, the *LX* line as a result of the selection reached a mean weight of 40 g. over its last six generations, and represents a considerable improvement over the original lines at their limits. Its progress is summarized in Fig. 2. The weights of the largest of the original lines (*CL*), over approximately the same period, are also shown in Fig. 2 for comparison. *CL* had reached a limit at 32 g., and the *LX* line eventually yielded an increase of 25% over this limit. Even compared to a later increase in the *CL* line, most likely due to a recombinational event (Roberts, 1966*b*), the *LX* line still shows a substantial improvement which must be attributed to the infusion of genes from the other selected large lines. In empirical terms, the *LX* line indicates clearly that crosses between the original lines at their limits yielded sufficient genetic variance for an appreciable further advance under selection.

The details of the response, however, are less clear. Some 18 months after the base population had been formed, it was by no means obvious from the 6-week weights of the 6th generation that any progress had been made. A promisingly high weight at the 4th generation had vanished as mysteriously as it had appeared. But after the 6th generation, there was a good response until a steady phase was reached by about the 13th generation. Though a linear fit would probably be an adequate description of the response retrospectively, a very different impression was formed as the data were collected. It seemed as if the response could be divided into three phases—an initial lag, a rapid response period, and a final limit. If this is so, then it is not at all typical of the asymptotic response curve classically expected of a selection programme. The initial lag differs also from what Mather & Harrison (1949) called 'delayed responses', which occurred after long periods of stability under selection.

If the suggested sigmoid shape of the response curve is real, one factor which

could explain it is linkage. If alleles that differed between lines were at loci that were linked, they would of course appear predominantly in the repulsion phase during the early generations, and progress under selection would depend on a sufficient number of cross-overs becoming available. If the postulated linkage were tight, this process would take a little time, though some progress would be expected from the start. Somewhat fortuitously, a partial check of the linkage hypothesis

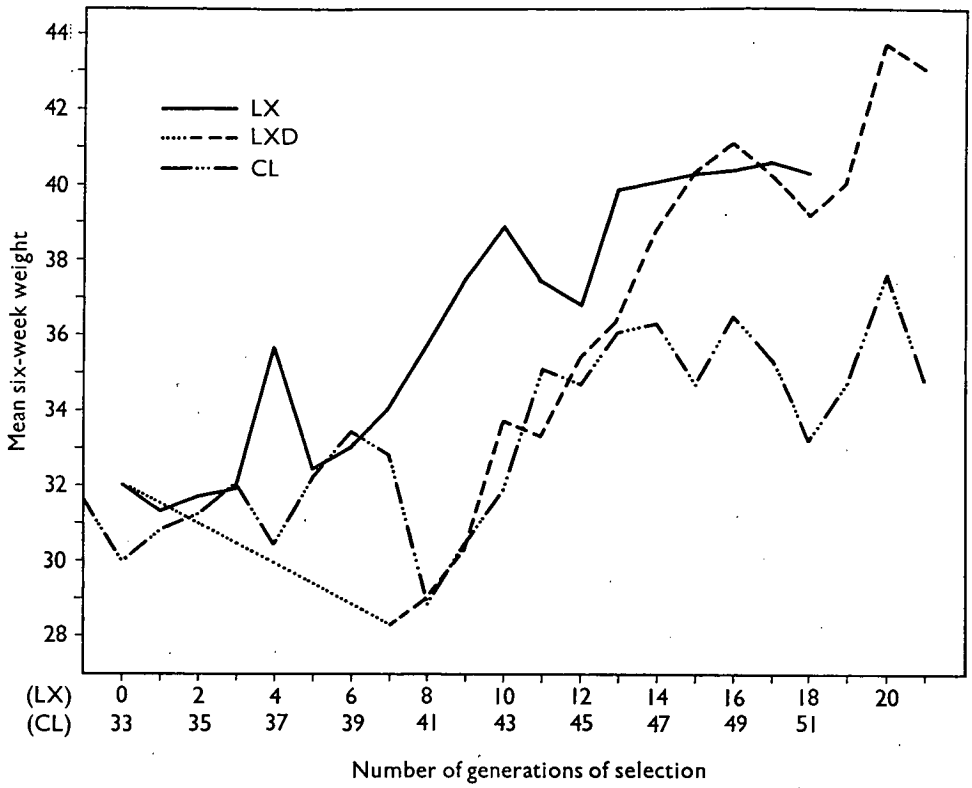


Fig. 2. Response to selection of *LX* and *LXD* lines. *CL* line shown for comparison.

was available when, by the 10th generation, the results suggested some such phenomenon. Thirty-five surplus litters from the LX_0 generation had been acquired by Dr Joyce Bloom for lung-tumor studies, described by Bloom (1964) and Falconer & Bloom (1962, 1964). From these litters, a control stock had been formed, which in the meantime had undergone six generations of random mating. Dr Bloom kindly allowed me to recover fifteen pairs of mice from different litters of her control stock, and these animals were mated appropriately to give a 7th generation of random mating. This formed a base population from which a second line, *LXD* (*D* for 'duplicate'), was selected for high 6-week weight. The mean weight of the base population of the *LXD* line is marked opposite the 7th generation of *LX* in Fig. 2. It can be seen that the random mating, or relaxed selection, had resulted in a drop

of about 4 g. since the zero generation. This, however, for present purposes, is inconsequential. The hypothesis to be tested was that, if linkage had impeded initial progress in the *LX* line, then the random mating ought to have allowed such linkage to break up, and that therefore the *LXD* line ought to give an immediate response when selection was applied to it.

The results, summarized also in Fig. 2, are easily compatible with this hypothesis. The response was indeed immediate, and despite its lower starting point, the *LXD* caught up with *LX* after eight generations of further selection. The relative rates of responses are seen more clearly in Fig. 3, which shows the generation means plotted against cumulated selection differentials. Over the period of the response, the

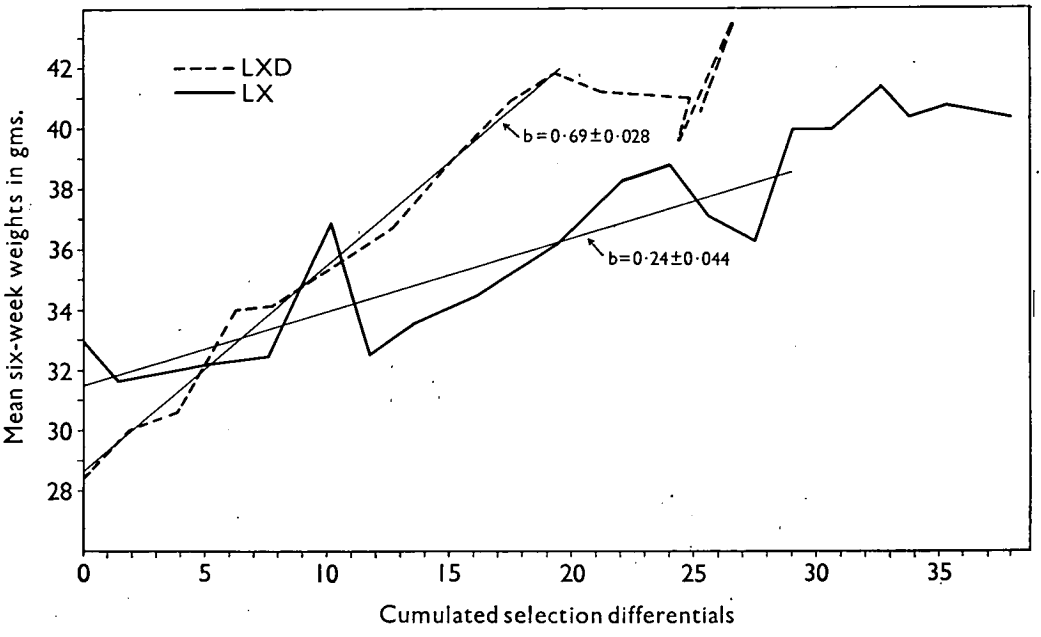


Fig. 3. Realized heritabilities in *LX* and *LXD* lines.

regression of generation means on cumulated selection differential, measuring the realized heritability, is much greater in *LXD* than in *LX*, the difference being significant beyond the 0·1% level (see Fig. 3). While this does not conclusively establish the linkage of genes affecting body weight in this population, no other explanation satisfies the facts with anything like the same facility.

If linkage did impede progress initially, then it is also possible that segments of chromosomes might have been fixed in the *LX* population while they were still in the repulsion phase, *i.e.* that some less favourable alleles, among those initially available, might have been fixed. This might be especially true of favourable alleles that had an initial frequency of 0·25, or of alleles linked to another with a greater effect on the character. If such were the case, then selection following a period of random mating might be expected to yield a further total advance under selection

than when the selection was applied from the start. The last two points of *LXD* are considerably higher than the final level of *LX*, suggesting that it had been advantageous to allow linkage to break up before selection was applied. However, these two points are based on the means of animals drawn from only seven and four litters, respectively, so that not a great deal of reliance can be placed on them. The *LXD* line (like *LX* before it) was approaching extinction through infertility by this time. Though tantalizingly suggestive, the results are therefore inconclusive on the question whether selection from crosses should be preceded by a period of random mating. In terms of applications to animal breeding, this is an important question which would merit further experimental investigation. For unless a greater advance is ultimately obtained, it is obviously inadvisable to delay the response by deliberately avoiding selection. An additional reason why selection should not be delayed for too long is that the *LXD* line regressed during the period of random mating. Though it obviously had not happened in this case, this could have meant that some alleles, or combinations of alleles, favouring large size might eventually have been eliminated from the population by natural selection acting against them.

The conclusions from this section are therefore that the original four large lines, at the limit, each lacked some genes contributing to large body size that were contained in one or more of the other three lines. It is also suggested strongly that, when the original lines were crossed, favourable alleles from different lines were put in the repulsion phase of linkage, and that this impeded the initial rate of advance if not the final limit.

(iii) *Further selection for small size*

The population, *SX*, formed by crossing three small lines at their limits, was subjected to continued selection for low 6-week weight. The results of this selection are summarized in Fig. 4. For comparison, the weights over the period of study of the *CS* line, the largest of the original small strains, are also shown in the figure.

For a long time, certainly up to generation 15, there was little if any evidence that the *SX* line had responded to selection at all. Since then, it has become more apparent that some progress has been made, though much of this impression stems from the last two points. The linear regression of generation means on cumulated selection differential was $-0{\cdot}124 \pm 0{\cdot}035$, which constitutes evidence of a significant response, albeit small. It is fair to add that this slope was increased from $-0{\cdot}083$ by the addition of the two final points.

The final points of the *SX* and *CS* lines shown in Fig. 4 were roughly contemporaneous, so it can be seen that the mean weights of *SX* have been lower than those of *CS* for several generations. This response, however, is much less than expected; the *CS* line had reached a limit to selection at around 14 g. (Roberts, 1966*a*) when it was crossed to the two other small lines, whose limits were about 11 and 10 g. Since genes from these smaller lines were at a frequency of at least 0·33 in the *SX* population, there is no obvious reason why the low weights of the smaller lines should

not have been regained by selection. The fertility of the *SX* line was consistently good and there is no likelihood that these genes were lost through drift. But the possibility that the lower limits found in previous experiments should be transcended, or even recovered, appears to be remote.

To what, then, must we ascribe the relatively poor response of the *SX* line? In an earlier paper (Roberts, 1966*b*) it was argued that the limit to selection for low 6-week weight in the *CS* line could be attributed to the opposing effect of natural selection acting on viability. In the case of *SX*, such an argument does not seem to apply. Some 95% of all matings were fertile, and viability over the critical period

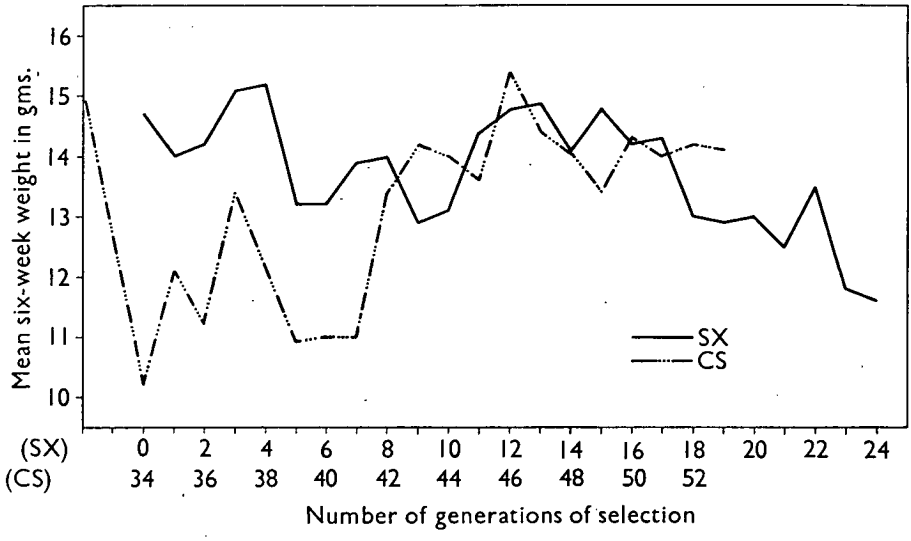


Fig. 4. Response to selection of *SX* line. *CS* line shown for comparison.

from birth to weaning was 96%, which is an extraordinarily good performance. Unless there was a great differential mortality of small mice *in utero*, it is difficult to see where any natural selection could have applied.

Another possible reason for the low response could be maternal effects. It is well known that in the mouse, a decrease in body weight leads to a reduced ovulation rate, and the mice gestated and reared in the consequently smaller litters have an advantage in body weight over mice from larger litters. (Selection for large size would equally lead to larger litters and a depressing effect on body weight.) However, these considerations, either, do not seem to apply to the *SX* line, as the mean litter size showed no evidence of any trend over the course of the experiment.

About the only remaining possibility is, again, linkage. It was seen from Fig. 1 that when the original small lines were crossed, body weight increased. This was not unexpected, for directional dominance is known to favour large size in the mouse. It does mean, however, that alleles for small size that differed in the three lines were put in repulsion on crossing, and that furthermore they would be masked

by the dominant alleles for larger size. If linkage is important, selection for small size would thus be expected to be ineffective in the early stages of the experiment. However, with crossing-over, coupling homozygotes should soon begin to appear, and selection for recessive genes, which is an efficient procedure, should yield a pronounced response once it began. This, obviously, did not happen, which means that if linkage is to be invoked as the full explanation of the poor response, we must stipulate that the linkage was very tight—far tighter than that which seemed to affect the loci controlling large size, discussed earlier. It did not even begin to break up until the 15th generation, and then only very slowly. To the extent that this is improbable, the linkage hypothesis lacks conviction as an adequate explanation of the slow response in the *SX* line. While linkage, almost certainly, impeded the response, it seems likely that it was augmented by some unidentified factor.

Whatever the full explanation may be, experience with the *SX* line provides a clear warning for those animal breeders concerned with the preservation of genes from declining breeds of livestock. It constitutes a strong empirical argument against tipping all these breeds into one gene pool. Even though desirable alleles are not lost through drift, they may not be easily recoverable from the pool—at least, not without more generations of selection than any breeder of large animals could cheerfully contemplate.

4. DISCUSSION

In as much as the crossing of selected strains generated new genetic variance and led to further responses to selection the results described above lend qualitative support to Falconer & King's (1953) procedure, quoted earlier. But the details are quite different. The heterosis found by Falconer & King when they crossed their two lines was only 5%; in the crosses reported here, the heterosis ranged from 8% to 32%, with an average of 16%. But the important difference is that Falconer & King did not find it necessary to suggest that progress had been impeded by linkage in their crossbred population. However, when their results are re-examined, the possibility of linkage cannot entirely be discounted. From their cross of the two large lines, they selected further for both high and low 6-week weight. After two generations by which time the cumulated selection differential was about 8 g. for the divergence, the high and low lines had failed to separate. The low line then came down, but their high line did not increase at all for another two generations. If we were now to wish to interpret these results in terms of linkage, we should obviously have little difficulty in doing so.

The relatively good response found by Falconer & King for downward selection from a cross of large lines has no bearing on the poor response reported here for the *SX* population, which was a cross of small lines. The two studies represent quite different situations.

The interpretation of the responses reported in this paper leans heavily on the hypothesis that linkage of loci affecting body weight was a prominent feature of the

crossbred base populations. There is, of course, no novelty in this suggestion. The influence of linkage on polygenic systems has long been discussed by Mather (see, for instance, his review, 1943), and a particularly clear case of linkage affecting sternopleural chaetae number in *Drosophila* was analysed by Thoday, Gibson & Spickett (1964). However, linkage in *Drosophila* is one thing; it would not necessarily lead one to expect the same phenomenon in an organism like the mouse, with twenty pairs of chromosomes. Now, the total number of genes, or effective factors, affecting body weight in the mouse is also of this order of magnitude (Roberts, 1966*a*). If these genes are linked to any important extent, it must mean that there is a considerable concentration of similar genes in certain segments of a few chromosomes. The phenomenon of clustering of functionally related genes is now well known in certain micro-organisms, although even among bacteria, it is by no means universal (Fargie & Holloway, 1965). As the clustering appears to be more widespread in bacteria than in higher organisms (Bodmer & Parsons, 1962), it would be most unexpected if close linkage were a basic feature of loci controlling body weight —a trait composed of diverse components—in the mouse. The linkage found, or suggested, in the experiments reported in this paper is much more likely to be the product of a special situation, as follows.

It is shown by Hill & Robertson (1966) that linkage affects the chance of fixation of alleles under selection. An unfavourable allele at a locus is more likely to become fixed if it is linked to another locus with a greater effect on the character. If the effects of the two loci are approximately equal, the chance of fixation of the more favourable allele is reduced at both loci. All this occurs even if the initial population is in linkage equilibrium.

Now, turning the argument around, this would suggest that under certain conditions, the only loci where an unfavourable allele is fixed are those that are linked to other loci affecting the character under selection. Loci that are unlinked would all be fixed for the more favourable allele, given those conditions. The conditions are the ones that exclude chance fixation, spelled out by Robertson (1960) and discussed by Roberts (1966*a*), who showed that these same conditions applied to all of the seven selected lines employed to form base populations for the studies described here. Therefore, when these lines were crossed, loci that were linked had sometimes been fixed for unfavourable alleles; and where the loci were of roughly equal effects, the allele fixed at a particular locus need not be the same for all the lines. Unlinked loci on the other hand, were largely fixed for the same alleles; the probability of this occurring was enhanced by some overlap in the origins of the various lines, as mentioned in an earlier paper (Roberts, 1966*a*). Genetic variance in the two crossbred populations would therefore be dominated by linked loci; unlinked loci would tend not to segregate and therefore contribute no variance.

If all this is correct, then the apparent importance of linkage in the *LX* and *SX* populations is largely an artefact of the method of construction of those populations. It does not necessarily mean that linkage generally affects the genetic variance of body weight in an unselected outbred population to anything like the same extent.

The relative importance of linkage will be the amount of genetic variance due to loci that are linked, as a proportion of the total variance in the character. It is suggested that this ratio is maximized in populations derived from crosses between lines that have previously been selected in the same direction.

SUMMARY

1. Four lines selected for large size were crossed to form a base population for further selection for high 6-week weight; three small lines were crossed similarly, and the crossbred population was selected for low 6-week weight.

2. In every case, a cross between two selected lines resulted in heterosis increasing body weight. This shows that all of the selected lines were differentiated with respect to genes affecting body weight.

3. Further selection for large size produced a stock whose mean weight was 25% higher than the largest of the original lines at its limit. But the response to selection for small size was slow, and after twenty-four generations of selection, the low weights of two of the original lines had not been recovered.

4. The evidence points to linkage of genes affecting body weight in the mouse. It is suggested that this is a particular feature of crosses between previously selected lines, rather than a general feature of mouse populations.

REFERENCES

BLOOM, J. L. (1964). Body size and lung-tumor susceptibility in outbred mice. *J. natn. Cancer Inst.* **33**, 509–606.

BODMER, W. F. & PARSONS, P. A. (1962). Linkage and recombination in evolution. *Adv. Genet.* **11**, 2–100.

FALCONER, D. S. (1960). Selection of mice for growth on high and low plains of nutrition. *Genet. Res.* **1**, 91–113.

FALCONER, D. S. (1964). Maternal effects and selection response. In *Genetics Today. Proc. XIth int. Congr. Genet. (The Hague)* 1963, Vol. III, 763–774.

FALCONER, D. S. & BLOOM, J. L. (1962). A genetic study of induced lung-tumors in mice. *Br. J. Cancer*, **16**, 665–685.

FALCONER, D. S. & BLOOM, J. L. (1964). Changes in susceptibility to urethane-induced lung tumors produced by selective breeding in mice. *Br. J. Cancer*, **18**, 322–332.

FALCONER, D. S. & KING, J. W. B. (1953). A study of selection limits in the mouse. *J. Genet.* **51**, 561–581.

FARGIE, B. & HOLLOWAY, B. W. (1965). Absence of clustering of functionally related genes in *Pseudomonas aeruginosa*. *Genet. Res.* **6**, 284–299.

GOODALE, H. D. (1938). A study of the inheritance of body weight in the albino mouse by selection. *J. Hered.* **29**, 101–112.

GOODALE, H. D. (1941). Progress report on possibilities in progeny-test breeding. *Science, N.Y.* **94**, 442–443.

HILL, W. G. & ROBERTSON, A. (1966). The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294.

MACARTHUR, J. W. (1944). Genetics of body size and related characters. I. Selecting small and large races of the laboratory mouse. *Am. Nat.* **78**, 142–157.

MACARTHUR, J. W. (1949). Selection for small and large size in the house mouse. *Genetics*, **34**, 194–209.

MATHER, K. (1943). Polygenic inheritance and natural selection. *Biol. Rev.* **18**, 32–64.

MATHER, K. & HARRISON, B. J. (1949). The manifold effect of selection. Part I. *Heredity, Lond.* **3**, 1–52.

ROBERTS, R. C. (1966*a*). The limits to artificial selection for body weight in the mouse. I. The limits attained in earlier experiments. *Genet. Res.* **8**, 347–360.

ROBERTS, R. C. (1966*b*). The limits to artificial selection for body weight in the mouse. II. The genetic nature of the limits. *Genet. Res.* **8**, 361–375.

ROBERTSON, A. (1960). A theory of limits in artificial selection. *Proc. R. Soc.* B, **153**, 234–249.

THODAY, J. M., GIBSON, J. B. & SPICKETT, S. G. (1964). Regular responses to selection. II. Recombination and accelerated response. *Genet. Res.* **5**, 1–19.

PAPER 8

The limits to artificial selection for body weight in the mouse. IV. Sources of new genetic variance - irradiation and out-crossing.

Genet. Res. Camb. 9, 87-98. 1967.

by

R.C. ROBERTS

Genet. Res., Camb. (1967), **9**, *pp.* 87–98
With 3 text-figures
Printed in Great Britain

The limits to artificial selection for body weight in the mouse

IV. SOURCES OF NEW GENETIC VARIANCE—IRRADIATION AND OUTCROSSING

By R. C. ROBERTS

A.R.C. Unit of Animal Genetics,
Institute of Animal Genetics, Edinburgh, 9

(*Received* 18 *August* 1966)

1. INTRODUCTION

In an earlier paper (Roberts, 1966*b*), it was established that a line of mice selected for high 6-week weight, after it had reached its limit to selection, contained no residual additive genetic variance in the trait. There was no regression of the weight of the offspring on that of the sire, and reversed selection failed to bring about any decrease in body weight. Though difficult to prove conclusively, it was suggested that most loci contributing to variance in body weight had been fixed in the population. This paper reports two attempts to introduce new genetic variance into the line, that would lead to a renewed response to selection for high 6-week weight. The first method attempted was irradiation; the second method was to outcross the selected line to a random-bred unselected population.

The induction of new genetic variance in quantitative characters by means of irradiation, more specifically X-irradiation, has been the subject of much comment. There are several references to the successful utilization of the method to improve varieties of commercial plants, though these successes may have required relatively high doses of irradiation—of the order of 100,000 r. Animal populations cannot be exposed to such high doses, and the results of similar studies on them are correspondingly less striking. Some early attempts to accelerate the advance under selection (for sternopleural bristles in *Drosophila*) by means of irradiation gave largely negative results (Serebrovsky, 1935; Rokizky, 1936). But a completely different outcome was reported by Scossiroli (1954). By alternating irradiation (3000 r. per generation) and selection, he increased spectacularly the number of sternopleural bristles in lines of *Drosophila* that had reached their limit to previous selection for a high score, although his low lines showed little further response under the same procedure. A similar experiment (Scossiroli & Scossiroli, 1959) using isogenic material gave essentially the same result. Clayton & Robertson (1955, 1964) employed a similar approach, but reported only modest gains, and were cautious about the general usefulness of the method, at least for mammals. Using 1800 r. per generation, they obtained small but consistent responses to selection for

sternital and sternopleural bristles from inbred lines of *Drosophila*. They calculated that 500,000 r. would be required to raise the genetic variance of an inbred line to the level of a standard outbred population. When the same dose of 1800 r. per generation was applied to seven selected lines that had reached their limits, the additional responses (above those of the non-irradiated controls) were likewise uniformly small, though significantly greater after irradiation in three of the seven lines.

Abplanalp, Lowry, Lerner & Dempster (1964) describe an attempt to exploit X-ray-induced mutations in subsequent selection for egg number in poultry. A total dose of 8000 r. was applied to chicken sperm over seven generations. Selection for a further six generations failed to show any improvement of the irradiated lines over their non-irradiated controls, and the authors conclude that the dose of 8000 r. did not induce sufficient new genetic variance to help selection for high egg number in chickens.

Though the evidence from the literature is inconsistent, it indicated that irradiation should be attempted as a method of inducing genetic variance in the line of mice that had ceased to respond to selection on account of the fixation of loci affecting body weight. In parallel with this study, a second method was explored. The large line was outcrossed to an unselected population to test whether any alleles were available in that population which were more favourable than the ones fixed in the large line. Some encouraging results from this method have been obtained recently by Robertson & Osman (private communication). They outcrossed to the base population a line of *Drosophila* that had reached its limit for low number of sternopleural bristles, selecting from the outcross to see how soon the original limit might be transcended, and by how much. Though, on average, their extra gains were small, some of their replicates surpassed the original limit by a substantial margin.

The *Drosophila* experiments suggest that both methods of introducing new genetic variance—by irradiation and by outcrossing—may result in some advance under further selection to a point beyond the previous limit. But the results are variable, and there appears to be little information about the general utility of the two methods for any mammal. The experiments described below were designed to explore their potential, using a line of mice that had reached its limit, and to formulate more clearly any problems that may have to be overcome when the methods are applied to a mammalian population.

2. THE FIRST EXPERIMENT—IRRADIATION

(i) *Materials and Methods*

The line of mice employed for this study was the *CL* line, described by Roberts (1966*b*). In generation 35, fifteen pairs of sibs were selected on the basis of their 6-week weights. The heavier and lighter member of each pair were assigned alternately into two matching groups. One group of fifteen males was then employed for continuation of the *CL* line while the other fifteen males were irradiated. After

ensuring that both testes were in the scrotum, the body was shielded except for the scrotal region, and a dose of 600 r. was applied to each male.

This dose produces many chromosomal aberrations in irradiated spermatids and later stages in the mouse. The males remain fertile for a short time and then they enter a sterile period. Six weeks or so after the irradiation, the males recover fertility, the mature sperm having been in the spermatogonial stage when irradiated. The gametes no longer contain chromosomal aberrations but they are expected to carry mutations at a frequency perhaps fifteen times higher than that in non-irradiated males. Female mice can hardly be irradiated at all without inducing complete sterility; even a dose as low as 50 r. destroys all early oocytes. The literature on the effects of irradiation on mammalian (especially mouse) germ cells is voluminous. A useful summary of the subject, as it governed the choice of procedures for this experiment, is provided by Russell, Russell & Oakberg (1958).

In view of the sterile period following irradiation, the fifteen males were not used for 3 months. One died during this time, but the surviving fourteen were then mated to females drawn from the succeeding (36th)generation of *CL*. These females were drawn in a manner identical to that described for the irradiated males. The mating was at random except for the avoidance of close relatives. A replicate of the *CL* line was therefore produced, the only difference between it and the parent line being the irradiation of the males with 600 r. This was the zero generation for further selection in an attempt to exploit any favourable mutations induced by the irradiation.

The mechanics of the selection programme to be followed will depend on whether the new mutations to be exploited (if favourable) are recessive or not. If they are dominant or semi-dominant, there is no problem; they will contribute to the variance in body weight of the progeny of the zero generation and will be selected in the normal course of events. But if they are recessive, steps must be taken to make them homozygous before they can be selected, and difficulties arise. Each mutation will presumably be a unique event and appear in only one gamete of the irradiated males. They will therefore be borne by single animals, and in the heterozygous state, in the first generation. However, after these heterozygotes have bred, half of their offspring (i.e. generation 2) will be expected to carry the particular mutation in which we may be interested. If these offspring are now sib mated, a quarter of all matings ought to be between animals heterozygous for the same original mutant; bearing in mind that in the grand parental generation there were two irradiated males, a quarter of the matings also could be between animals heterozygous for a mutant from the second male, should that male as well have passed on an autosomal mutant to one of its progeny. Uncertainties about the number of loci involved, and the number of mutations an irradiated gamete may be expected to carry, rule out any probability statements about the frequency of homozygotes to be expected following sib mating, but subjectively, the procedure was thought to be worth the attempt.

Hence, at generation 2, the irradiated line was split. One subline was designated *ID* (*I* for 'irradiation' and *D* for 'dominants'); this line aimed to exploit dominant

and semi-dominant mutations at loci affecting body weight, and was selected and mated in the same way as the parent *CL* line. The other subline was designated *IR* (*R* for 'recessives') and the generation consisted of matings between sibs, for the reasons given above. The progeny of the sib matings were then selected, hopefully to pick out any desired recessive homozygotes, and then mated according to the usual scheme of avoiding close relatives. The sib mating, with selection, was repeated twice, separated by a generation of selection and mating without inbreeding. Thereafter the *IR* line was treated in the same way as *ID*.

Both lines, and also the parent *CL* line were run on fifteen pair matings. Within-family selection was practised throughout, to avoid complications due to maternal effects in the interpretation of the results.

The irradiation was applied once only, in the beginning. In retrospect, this was perhaps a mistake, but at the time, a single acute dose was deemed sufficient to test the general utility of the method. A purely operational point of view was taken—'Does 600 r. give us anything worth while?'. With the power of hindsight, a better question might have been—'What total dose is required to give us anything at all?'

(ii) *Results and Discussion*

The main results from the irradiation experiment are presented in Fig. 1. This shows the mean 6-week weights of the two irradiated lines compared to those of the parent *CL* line (the dotted parts of the *IR* graph refer to the sib-matings, mentioned above). The interpretation of the results is complicated slightly by an increase in weight in the *CL* line at the 44th generation. This increase was discussed fully by Roberts (1966*b*) and attributed to a genetic change in the line, most probably a rare recombinational event, though some increase due to environmental causes could not be completely discounted. Figure 1 now confirms that there was no marked increase from a general environmental influence at the time; the *IR* and *ID* lines failed to show any parallel increase, the generations in vertical alignment in the figure being roughly contemporaneous in all cases. We should therefore judge whether there has been any additional response to selection in the *IR* and *ID* lines, following the irradiation, by comparing them with the old level of *CL*, i.e. generation 43 and earlier. The limit prior to this point was established as 32 g., whereas subsequently it rose to 35 g. (Roberts, 1966*b*).

Nine generations after the irradiation, there was no evidence that selection in the *ID* and *IR* lines had increased weights at all over the original level of 32 g. Then both lines showed some increase. The two irradiated lines by this time were running so obviously in parallel that the *ID* line was discontinued after generation 12. The remaining line, *IR*, settled down to a level on average about a gramme or so higher than the original limit in the *CL* line.

The question then arises whether this rather small increase was attributable to a response following the irradiation. Quite obviously, it might have been, but it is hard to say for certain. Changes in the weights of selected lines of mice at their

limits are common, as seen from an earlier paper in this series (Roberts, 1966*a*). In any other circumstances, a shift of this magnitude in mean weight would attract little attention, for environmental trends in long-term experiments may always be suspected. But if we were prepared to dismiss this possibility, the following points could be listed in favour of the possible effects of the irradiation:

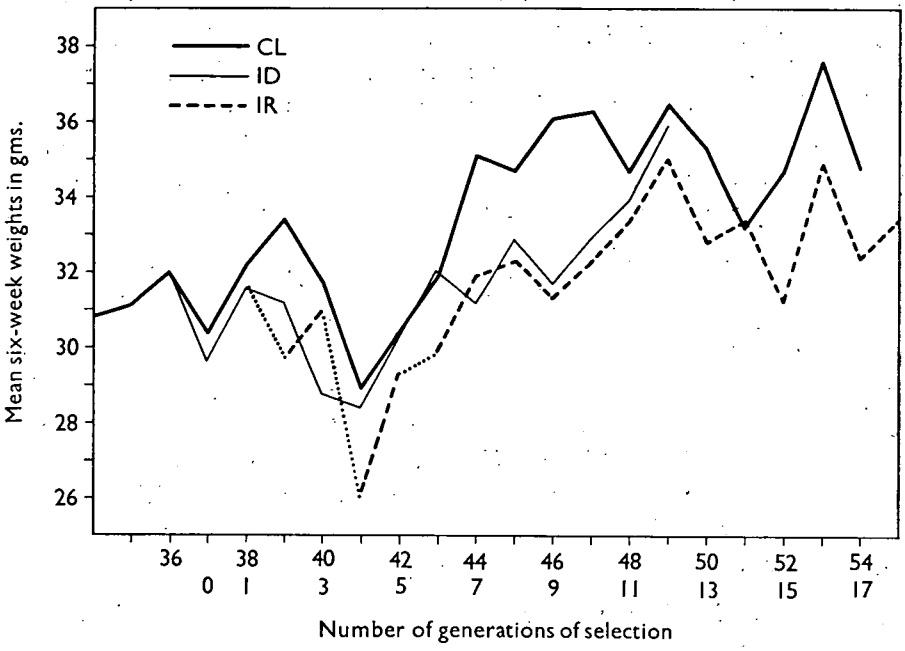


Fig. 1. Responses to selection for body weight following irradiation in *ID* and *IR* lines. Non-irradiated parent line (*CL*) shown for comparison. The dotted parts of *IR* line correspond to generations of sib-mating (see text).

(1) There was a slight increase in absolute terms, as mentioned.

(2) The postulated recombinational event that increased the weight of *CL* must be so rare that it cannot be invoked a second time; nor should any other rare recombination occur synchronously in both irradiated lines.

(3) There was some suspicion of an increase in variance within sexes within litters—the variance upon which the selection acted—following the irradiation. This increase over the *CL* variance, however, was by no means consistently a feature of the irradiated lines over succeeding generations.

(4) In one of the irradiated lines (*IR*) the regression of offspring weight on that of the sire became significantly positive ($+0{\cdot}133 \pm 0{\cdot}059$). This contrasts sharply with the position in the *CL* line and some of its other derivatives (Roberts, 1966*b*). Counterbalancing this argument, the same regression in *ID* was much less than its standard error, to which we may add the fact that in the remainder of the material for this series of papers, a sufficient number of

insignificant regressions of this type have been calculated that an apparently significant one, by chance alone, should be expected.

(5) Possibly the strongest argument in favour of an effect of irradiation is that, if there was any effect, the increase in weight found is very much what would be expected, if Clayton & Robertson's (1964) *Drosophila* bristles provide any lead.

To summarize, the main conclusion is that a gonad dose of 600 r. to male mice, from a selected line at its limit, did not contribute to any substantial advance under further selection. But there may have been *some* advance. If an experiment were planned to investigate this possibility further, the following recommendations could now be made:

(1) A single dose of 600 r. is too small, in the light of general experience with other organisms. Of particular interest in this context is Russell's (1962) finding that X-rays delivered in two fractions of 500 r., separated by 24 hours, gives a mutation rate of five times that observed for a single 1000 r. dose given to male mice.

(2) As success, if any, may be sporadic and unpredictable, a proper experiment on irradiation effects should incorporate several replicates.

(3) The control lines should also be replicated, to safeguard against fortuitous shifts of the kind found in the *CL* line.

(4) To allow clear patterns to emerge, the considerable facilities that all this involves may have to be committed to the project for at least twenty generations.

As a postscript, in view of the general interest in the effects of irradiation on

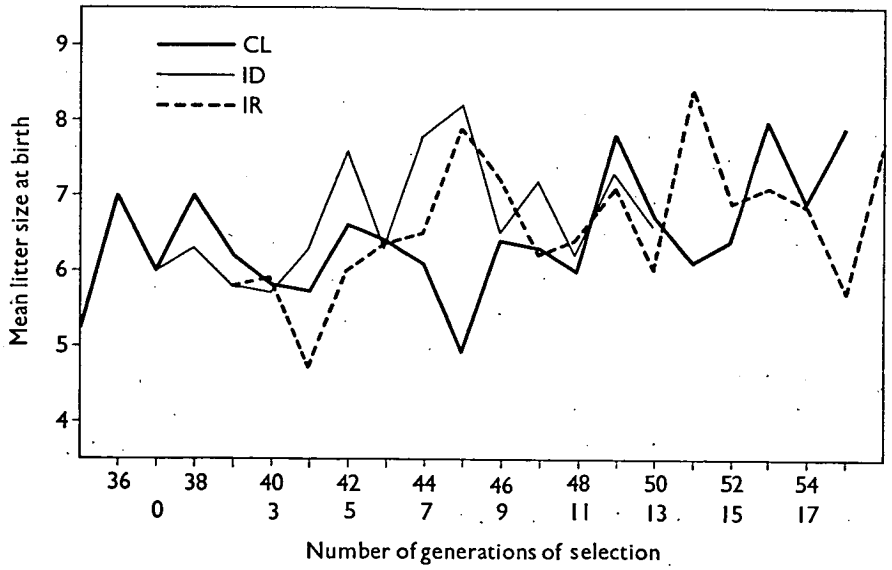


Fig. 2. Effects of irradiation on litter size in *ID* and *IR* lines. Non-irradiated parent line (*CL*) shown for comparison.

fertility, the mean numbers of live young in the first litters of the three lines are shown in Fig. 2. Litter size, while it may not have increased much following the irradiation, certainly did not show the expected decline.

3. THE SECOND EXPERIMENT—OUTCROSSING

(i) *Materials and Methods*

It was explained earlier that the purpose here was to comb a random-bred population for any alleles more favourable to high weight than those fixed in the *CL* line. A potentially useful source of such alleles was locally available in Dr D. S. Falconer's *Q* strain, whose origin and structure is described in *Mouse News Letter* (No. 25, p. 29). The strain was constructed from a broad base, though one which had considerable overlap with the *CL* line.

The procedure for drawing two matching samples from the *CL* line was described in the previous section, when choosing fifteen males to be irradiated. In an identical manner, fifteen females were drawn from the same (35th) generation of *CL*. These were mated to fifteen males from the *Q* strain to form the zero generation of the *CQ* line, from which to select for an increase in 6-week weight. Two questions were asked of the selection programme:

(1) How long would it take to restore the mean weight to the level of *CL*? There was no reason why this level should not be regained, since genes that had been fixed in *CL* were at a frequency of at least 50% in the *CQ* population. If alleles from *CL* were superior at all loci to those available in *Q*, it would merely be a question of making them homozygous again to restore the level of the *CL* body weights.

(2) Could the limit reached in *CL* be transcended, and if so, by how much? This is equivalent to asking whether any of the *Q* alleles were superior to those fixed in *CL*.

Depending on the effect of the infusion of genes from the *Q* strain, it was anticipated that some selection would be required to nullify the effects of the outcross on body weight. It might therefore be advantageous to start the new selection from a higher level. To test this, the matching procedure was employed again to divide females of the zero generation of *CQ* into two groups. One group was mated to males selected from the same generation to continue the *CQ* line, while the other group of females was backcrossed to males of the 36th generation of *CL*, which were again drawn alternately from selected sib pairs. This second set of matings was designated the zero generation of the *CQB* line, which from that time was treated exactly the same as the *CQ* line.

It was expected that *CQB* might regain the original limit of *CL* sooner than *CQ*, but that this might involve some sacrifice of the total advance. The reason for this would be that any favourable alleles from *Q* that were at a low frequency might be lost during the backcrossing, by sampling. But if favourable alleles were at a high

frequency in the *Q* stock, the backcrossing should not affect greatly their availability for further selection.

The approach taken at the start of these studies was quite empirical, and the results are presented here as such. However, the implications of the procedure and its theoretical basis will be described in more depth in a forthcoming paper by Robertson & Osman (private communication), who carried out a similar but more extensive study on *Drosophila*.

As in other lines described in this series of papers, *CQ* and *CQB* were both run on fifteen pair matings per generation, and the within-family method of selection was employed throughout.

(ii) *Results and Discussion*

The mean 6-week weight of the *Q* strain when it was crossed to *CL* was about 22 g. The mean weight of the F_1 was above the mid-parental value, as expected. This is because directional dominance is towards large size, and also because of the maternal effect deriving from the use of the large *CL* mice as the dams during the crossing.

The progress of the *CQ* line from zero generation is shown in Fig. 3. After six generations of selection, the weights were actually lower than at the starting point. There was then a steady response for perhaps ten generations, after which weights fluctuated erratically around a mean of 36 g. or so.

The *CQB* line shows a very similar pattern of response. On backcrossing (using *CL* males on *CQ* females), the weights were actually lower than those obtained in the *CQ* line (*CQ* females by *CQ* males). This was probably an accident of sampling, but over the succeeding generations there is little evidence that the weights had been increased by the backcross over the level of the outcross. Again, as in *CQ*, the *CQB* line showed no advance under further selection for six generations. This was followed by a sharp response, and by the 12th generation, it appeared as if *CQB* was going to exceed the level attained by the *CQ* line. But eventually both lines settled down to much the same weight, at around 36 g.

The main interest in this study concerns the final limit of the *CQ* and *CQB* lines compared to the initial limit of 32 g. in the *CL* line. (The complication of a later rise in *CL* is disregarded for reasons given earlier.) There is no doubt that genes have been extracted from the *Q* strain that enable the limit of the *CL* line to be surpassed. This means that less favourable alleles had been fixed at some loci in the *CL* line during the earlier selection.

However, the *Q* strain has proved less useful as a source of new variance in the *CL* line than did three selected large lines, in a study described by Roberts (1967). There the gain ultimately attained over the *CL* level was fully twice that reported here.

A point of considerable practical importance is the length of time required after the outcross to recover the weight of the selected line. Again taking 32 g. as the original level of the *CL* line, this weight was reached by the 9th generation in both *CQ* and *CQB*. As *CQB* had required an extra generation for the backcross, it is

therefore marginally inferior on this score to the *CQ* line that was selected straight away from the outcross. The idea that it would be advantageous to start from a higher level, and recover the lost weight sooner, was thus not sustained in practice.

The lag of six generations after crossing before any response to selection was observed, is an exact repeat of the experience from selection following the crossing of selected strains (Roberts, 1967). The results from that study were interpreted in terms of linkage, and without repeating any of the argument, the same explanation is equally satisfactory here.

The close similarity between the *CQ* and *CQB* responses is quite striking. Generation n of *CQB* was always roughly contemporaneous with generation $(n+1)$ of *CQ*. This dispels the possibility that environmental trends affected the two lines synchronously. For instance, when *CQ* began to respond, *CQB* adopted the same

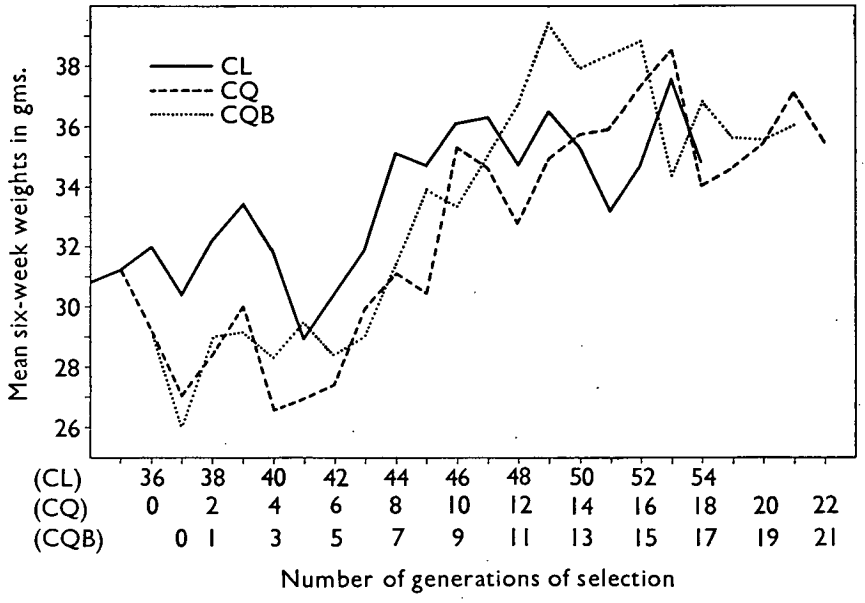


Fig. 3. Selection for body weight from cross between a line at its limit (*CL*) to a random-bred strain. *CQ*-selection from F_1. *CQB*-selection from backcross to *CL*.

pattern 3 months later. The similarity suggests that the *CQB* line was not adversely affected by the fact that new genes from *Q* had their frequencies halved by the backcrossing. This must mean that favourable alleles from *Q* were not at a low frequency in that strain; had they been rare, it would have taken a little while longer in *CQB* to increase their frequency, before their effects on the mean weight became noticeable.

Outcrossing, as a method of breaking through the limit, therefore seems to have been moderately successful. This encourages the thought that other outcrosses to unrelated populations might bring about further gains. This idea, of course, is not

new. Falconer (1960), discussing the relatively small divergences generated by laboratory selection experiments, compared to differences between breeds of livestock, wrote: 'The reason for the disappointing results of experimental selection ... is that experiments are carried out with closed populations of not very large size. The limits are set by the gene content of the foundation individuals.... The breeder of domestic animals, in contrast, by intermittent crossing casts his net far wider in the search for genes favourable to his purposes.'

4. GENERAL CONCLUSIONS AND IMPLICATIONS

Irradiation, as a means of generating new genetic variance, does not inspire much hope in terms of the improvement of mammalian populations. It is not easy to apply in practice, and the expected gains, if any, are small. If the study reported here is typical of what may be found, special attempts to extract favourable recessive mutations do not lead to any increased gains. This statement, however, may not have general validity. It is probably true of characters like large size, where favourable genes tend to be dominant. But the desired expression of other characters may involve recessive genes, and the prospects of their recovery from irradiated material is not good. For these traits, especially, irradiation as a method seems to have little to commend in it.

Outcrossing to an unselected population seems more hopeful, but it may take eight to ten generations to recover the level of the original limit. For most farm livestock, this is a discouraging prospect, and one that could not be undertaken by individual breeders. The time scale for most domestic animals would be 10 to 20 years, before any increased return might be expected. Only then could any of the lost production in the interim begin to be recovered. Even on the basis of a national scheme, stock improvement by outcrossing to genetically inferior material seems feasible only for rapidly reproducing species, where the breeding project need not encroach heavily on the current production facilities.

The greatest improvement over the initial limits, found in the investigations reported in this series of papers, came from crosses between selected strains with further selection from the crosses (Roberts, 1967). Even this method, though highly successful in increasing body weight, was not very effective in reducing body weight. It may therefore not always work in the desired direction. Where it was effective, it became associated with fertility problems that demanded special attention. But its main advantage over the outcrossing method was that there was no regression from the level of the initial limit, at least not with crosses between large strains. If small body weight were desired (as, for instance, in egg-laying strains of poultry) the increase in weight on crossing may detract from the value of the method, even without the doubt about the effectiveness of further selection for small size.

Linkage seemed to impede progress under further selection in all material involving crosses. It was suggested in the previous paper that this problem may be common in crosses involving highly selected lines or strains, since the unfavourable

alleles likely to be fixed are those linked to others affecting the trait. Linkage may reduce somewhat the further gains that may be made, but its chief nuisance value is in impeding the initial rate of advance, if the experiments reported in this and previous papers are representative.

In summary, the limits to artificial selection, being a function of the gene content of the selected material, need not be insuperable barriers to further progress if a useful source of better genes can be tapped. But with slow-reproducing mammals, especially, any method employed to transcend the limit is likely to be time-consuming and costly. The problems are those of organization and finance; the genetic methods have been examined and, within their context, evaluated in this series of papers.

SUMMARY

1. Two methods are examined of introducing new genetic variance into a line of mice selected for high 6-week weight which, at its limit, displayed no additive genetic variance.

2. The first method—irradiation—gave largely negative results. Any further gain under selection that was achieved could not be clearly distinguished from a possible environmental trend.

3. The second method—outcrossing to an unselected strain and then selecting from the cross—resulted in a clear gain over the original limit, but nine generations were required even to recover the original limit.

4. Various methods of transcending selection limits are evaluated in terms of their application to livestock improvement.

I am greatly indebted to Mr J. H. Isaacson for irradiating the mice for me.

REFERENCES

ABPLANALP, H., LOWRY, D. C., LERNER, I. M. & DEMPSTER, E. R. (1964). Selection for egg number with X-ray-induced variation. *Genetics*, **50**, 1083–1100.

CLAYTON, G. & ROBERTSON, A. (1955). Mutation and quantitative variation. *Am. Nat.* **89**, 151–158.

CLAYTON, G. & ROBERTSON, A. (1964). The effects of X-rays on quantitative characters. *Genet. Res.* **5**, 410–422.

FALCONER, D. S. (1960). *Introduction to quantitative genetics*, pp. 216–217. Edinburgh and London: Oliver and Boyd.

ROBERTS, R. C. (1966*a*). The limits to artificial selection for body weight in the mouse. I. The limits attained in earlier experiments. *Genet. Res.* **8**, 347–360.

ROBERTS, R. C. (1966*b*). The limits to artificial selection for body weight in the mouse. II. The genetic nature of the limits. *Genet. Res.* **8**, 361–375.

ROBERTS, R. C. (1967). The limits to artificial selection for body weight in the mouse. III. Selection from crosses between previously selected lines. *Genet. Res.* **9**, 87–98.

ROBERTSON, A. & OSMAN, H. E. (private communication).

ROKIZKY, P. (1936). Experimental analysis of the problems of selection by X-ray irradiation. *Usp. zootekh. Nauk*, **2**, 161–202.

RUSSELL, W. L. (1962). An augmenting effect of dose fractionation on radiation-induced mutation rate in mice. *Proc. natn. Acad. Sci. U.S.A.* **48**, 1724–1727.

RUSSELL, W. L., RUSSELL, L. B. & OAKBERG, E. F. (1958). Radiation genetics of mammals. Ch. 8 (pp. 189–205) in *Radiation biology and medicine* (W. D. Claus, ed.). Reading, Mass.: Addison-Wesley Publishing Co. Inc.

SCOSSIROLI, R. E. (1954). Effectiveness of artificial selection under irradiation of plateaued populations of *Drosophila melanogaster*. *Symp. Genetics of Population Structure, Pavia* (A. A. Buzzati-Traverso, ed.), pp. 42–66.

SCOSSIROLI, R. E. & SCOSSIROLI, S. (1959). On the relative role of mutation and recombination in responses to selection for polygenic traits in irradiated populations of *Drosophila melanogaster*. *Int. J. Radiat. Biol.* **1**, 61–69.

SEREBROVSKY, R. E. (1935). Acceleration of the rate of selection of quantitative characters in *D. melanogaster* by the action of X-rays. *Zool. Zh.* **14**, 465–480.

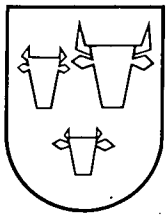
PAPER 9

Selection limits in the mouse and their relevance to animal breeding.

Proc. 1st World Cong. on Genet. appl. to Anim. Prod., Vol. 1, pp. 493-509, Madrid, 1974

by

R.C. ROBERTS.



Ist. WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION

Ier. CONGRES MONDIAL DE GENETIQUE APPLIQUEE A L'ELEVAGE

I CONGRESO MUNDIAL DE GENETICA APLICADA A LA PRODUCCION GANADERA

I. WELTKONGRESS UEBER ANGEWANDTE GENETIK IN DEN LANDWIRTSCHAFTLICHEN NUTZTIEREN

REPRINT

TIRE-A-PART

SEPARATA

MADRID

Spain - Espagne ● España - Spanien

7-11 Octubre - Oktober
October - Octobre 1974

Depósito legal: M. 28722.—1974.

Gráficas Orbe, S. L., Padilla, 82. Madrid.

SELECTION LIMITS IN THE MOUSE AND THEIR RELEVANCE TO ANIMAL BREEDING

Limites à la sélection chez la souris et ses implications dans l'élevage animal

Los límites a la selección en experiencias con ratones y sus implicaciones en la mejora animal

R. C. ROBERTS *

The pattern of response to artificial selection has a well-defined expectation: the response diminishes progressively until it reaches the asymptote, when the population is said to be at its limit to selection. Perhaps three aspects of this generalisation should be emphasised.

Firstly, the actual limit reached in any particular case depends on the environment in which the selection is conducted. It is obvious that selection for milk yield or for fat lamb production would not reach the same limit in poor upland conditions as it would in a more favoured geographical area. It is furthermore true that the limit may not be identical when populations selected in different areas are transferred to common ground. FALCONER (1952, 1960*a*) discussed how different genes may contribute to the response in rich and poor conditions. In other words, the environment defines the character, and clearly we have no *a priori* grounds for expecting the limit to be the same for what are, effectively, two characters. BATEMAN (1971), reporting on the first five generations of selection for growth in the mouse, conducted over a range of diets varying in the proportions of milk and maize, was already finding hints of special aptitudes developing on the more diverse regimes. In such a case, no meaningful statement about the limits to selection for growth could be made without specifying the diet under which the selection was done. The idea that some genes contribute to the response only in a specific environment further implies that once a population has been selected to the limit, it is in principle possible to find a new environment in which a further response might be obtained, but as far as I am aware, we have no experimental evidence to support this suggestion. A much fuller discussion

* Agricultural Research Council, Unit of Animal Genetics; Institute of Animal Genetics, West Mains Road, Edinburgh EH9 3JN, Scotland, U. K.

of the interactions between the genotype and the environment is given by DICKERSON (1955, 1963), where he introduces such terms as the «treadmill environment» and «genetic slippage».

The second aspect which affects our interpretation of selection limits is more directly genetical. A theoretical treatment of limits by ROBERTSON (1960) establishes that the limit reached is a function of Ni, where N is the effective population size and i is the intensity of selection. Thus, intense selection in a small population will lead to a different limit than will be the case when the selection is less intense, or where the population size may be larger. These factors are discussed in more detail, with particular reference to experiments with *Drosophila*, by OSMAN and ROBERTSON (1968). Different methods of selection, as is well known, affect both the intensity and effective number, and consequently, different methods may lead to different limits. Again, experimental evidence lags behind theory, and especially so in the case of mammals. Nevertheless, the responses reported by COMSTOCK (1973), for growth in mice, continued for a much longer period than they did in similar experiments conducted in Edinburgh (ROBERTS, 1966*a*, FALCONER, 1974). Comstock's total gain for increased growth amounted to 16-20 genetic standard deviations, whereas the Edinburgh experiments summarised by ROBERTS (1966*a*) yielded 4-12 of the same units. The relevant point in the present context is that Comstock's effective population sizes were 40 or more, while the Edinburgh populations were about half of this size. This serves to illustrate how the limit actually reached may be very much a function of the selection programme which generated it. It is, however, quite impossible to make detailed comparisons between sets of data from different places. The factor which will affect the limit much more than any other is the gene content of base population to which selection is applied.

The third general point concerning limits is that the classical expectation of a selection response, as stated at the outset, applies primarily to large populations, with a large number of segregating genes, whose individual effects are small, where changes in gene frequencies are untrammelled by the complications of inbreeding or drift, or the linkage of genes affecting the character. In other words, it refers to idealised situations. It is neither my purpose nor my mandate here to examine what are mainly theoretical considerations, but in attempting to review studies on limits conducted with the laboratory mouse, it may be important to remember that in some respects, the departures from the idealised situation may be quite marked. Population sizes in mouse work, as has been just noted, tend to be rather small; as will be discussed later, genes with large effects have been established with a fair degree of certainty; linkage of genes controlling a character have been invoked, while in one case, at least, non-additive genetic variance was found to be present to an extent not readily accommodated by theoretical considerations of what to expect. It was mentioned earlier that the theoretical treatment of limits is as yet inadequately supported by experimental data. It is also true that departures from the existing theoretical models generate some empiricism which rather suggest the need for more theory.

We can thus see that the limit attained in any particular selection experiment is a function of the environment in which it was produced, the genetic properties of the population in which it was obtained, and the details of the methods which led to it. These and other constraints emphasise that the results of experiments on

limits using the laboratory mouse may not be directly applicable to a national livestock improvement scheme, or to the operations of an international poultry company. Over and above these particular reservations, there are the general difficulties of translating across species. The mouse, for all the fact that it is a mammal, is not a transistorised pig, while it is easier to think of physiological functions that distinguish it from the hen than it is to exemplify functions which the two have in common. And then there is the question of whether the laboratory environment is a sufficiently good analogue of the field or cattle yard to yield results that are comparable in genetic terms. Leaving aside the asepticism of certain hyperclinical laboratories, which certainly put themselves *hors concours* in terms of many agricultural applications, it is still true that laboratory mice are typically kept in a very uniform environment. If a fluctuating environment leads to any particular demands on the genotype — and little seems to be known whether it does or not — the laboratory mouse would not seem to have much that is germane to say on the matter.

Some of these constraints would be less severe if selection limits could be regarded as observations or measurements in their own right, but they are nothing of the kind. They are the end product of a whole galaxy of factors surrounding particular selection programmes extending over many generations. When we talk of a selection limit, more so than in the case of most genetic phenomena, we are talking of a highly specific set of conditions, and the generality of the result should, as a basic minimum, be pondered.

Nevertheless, cry *caveat emptor* though we may, we cannot stop here. For one thing, it is basically unhelpful to establish complications and then leave them. For another, science would stop if every result were limited by its own specificity. Only as generalities emerge do principles become established, and selection limits are subjected to the same rules. Ideally, therefore, we should consider results from many species in many situations, to see how clever we may be at detecting patterns. Unfortunately, the material is not available for this exercise, and we are restricted to laboratory species. We should probably be glad, in view of the current demands on the livestock industry, that most farm animals are not yet at their limit to selection and gains are still to be made. The most notable exception to this statement is the egg-laying sector of poultry, though even there commercial breeders may have successfully counteracted some deterioration in the environment, through intensification and also disease. Selection for broiler chickens may already be facing diminishing gains (CLAYTON, 1972*a*) and consequently an incipient limit. It is possible that within the foreseeable future, some pig improvement schemes may also reach this stage. Whereas it is unlikely that cattle or sheep will present a similar problem for a long time to come, it is still the case that the breeders of some important classes of livestock should now be considering how they could aspire to further gains, once the current selection programmes run out of steam. And for obvious reasons, the consideration is therefore mostly restricted to a theoretical approach and the evidence from laboratory animals. It is almost certainly true that various commercial poultry enterprises have also addressed themselves to the problem, but their results are not public property. This review is confined to the laboratory mouse — the only mammalian species from which results are available.

The idea that limits were inevitable preceded any experimental demonstration

of their existence in any animal species. In his 1950 book on *Population Genetics and Animal Improvement,* Lerner states explicitly that «the limit of selection progress will be reached when all of the potential variability is converted into free form and subsequently exhausted by fixation». He goes on to add that «it is rather unlikely that this eventuality will arise before other factors limiting selection progress begin to operate». MATHER (1949) was also very much aware of the concept, though his approach was different. He was concerned with the assembly of positive allelomorphs among the descendants of crosses between plant varieties, and pointed to the difficulties that could be created by close repulsion linkage. Though MATHER seemed to regard limits at that time in a short-term predictive context, he obviously interpreted them very clearly in terms of the gene content of the base material.

As will be shown shortly, both MATHER'S and LERNER'S views had strong prognostic features, and were clearly much more realistic than GOODALE'S (1941) view, who saw selective breeding as preferable to «the preservation of haphazard but inheritable modifications, known as mutations», to study evolutionary changes. In his view, «nearby limits» were just temporary impediments, to be overcome by perseverance. GOODALE'S views are not quoted with any sense of disparagement, but rather the contrary; he was an eminent and much-respected pioneer of experimental work in quantitative genetics, whose contribution was immensely influential. The point is that, as might be expected, selection limits did not become an experimental problem until the experimentalists had experienced them, though those people well versed in selection theory quite clearly saw them coming. LUSH (1945), in his well-known book, at least implicitly deals with all the factors we know today that lead to limits. As a specific example, he discusses quite explicitly how selection against recessives becomes increasingly ineffective as the gene frequencies become low. This obviously implies the asymptotic approach to a limit through the fixation of segregating genes. Other examples from LUSH'S book, like his treatment of heterozygous advantage or opposing natural selection, would demonstrate that the mechanisms which lead to limits were perfectly well understood. It is simply that the emphasis was different: in the context of their time, these mechanisms were viewed as slowing down the response to selection, rather than stopping it altogether. The point has been made earlier in this article that the concept of a limit is not an autonomous one: it is the end point of a process, and it can be understood only as a function of the whole process.

The first experiments on selection limits in mice arose directly from the pioneering work of GOODALE (1938, 1941) and of MACARTHUR (1944, 1949). Both had in fact selected their lines to their limits for increased body weight, though neither seems to have made the claim explicitly. GOODALE, as noted earlier, was, shall we say, not looking for the phenomenon. MACARTHUR noted in his 1949 publication that his response was tailing off, and BUTLER (1952), working on the same material, found that little further progress had been made. In the late '40's, samples of both GOODALE'S and MACARTHUR'S large strains were obtained by FALCONER in Edinburgh, as reported by FALCONER and KING (1953). They confirmed that there was no response to further selection for high body weight, though a slight response to reversed selection, in both strains, indicated that fixation was not complete. FALCONER and KING noted that the two strains had, at least to some extent, achieved their high body weights for different reasons. GOODALE'S strain, they

observed, was large-bodied but not very fat, whereas MACARTHUR's strain was smaller in linear dimensions but was very fat. From this, they argued that each strain might have genes for large size which the other lacked, and that a cross between the two might therefore yield new genetic variance to render possible a further response to selection, beyond the original limits. This expectation was amply confirmed by the experiment they reported, and body weights were increased by about 10 % over nine generations of further selection. This was clear proof that the two large lines were differentiated genetically, and that the initial limits were to be attributed to the loss of additive genetic variance through the fixation of genes affecting body size. Further proof that new genetic variance had been generated was given by the results of selection for smaller body weights from the cross. This was conspicously more successful than the reversed selection from the two lines independently.

Another important early experiment on selection limits in the mouse was reported by LEWIS and WARWICK (1953). This is a much-neglected piece of work, at least in the context of limits. One possible reason for the neglect is that the word «limits» did not appear in the title of the paper, which may be a sad reflection on the way we treat the older literature. A further reason is that the authors themselves were more concerned with the effect of mating system on progress under selection, and on that score found a negative result. However, what they did in setting up the experiment had a much wider significance. They were working on MACARTHUR's strains, both large and small, and outcrossed each to an unselected stock from the same base population as the selected lines. After one backcross to each of the large and small lines, they continued to select for large and small body weight, respectively. Over five generations of selection, they obtained significant responses in both directions. In view of BUTLER's (1952) finding that the lines had by that time reached their limits, the conclusion is unambiguous that an infusion of genes from the base population had been responsible for the renewed response. No doubt in those days the reason would have been attributed to chance variation in the sampling of the base population, to set up the different stocks, and this factor could indeed have had some effect. But this was before the publication of KIMURA's (1957) much-used formula for the chance fixation of genes, and the subsequent development of this line of thinking indicates that LEWIS and WARWICK may have recovered some genes from the base population that had been lost, by chance, from MACARTHUR's lines during selection.

The model of selection limits posed by these early mouse experiments is therefore one of the exhaustion of the additive genetic variance through the fixation — perhaps partly through chance — of genes contributing to the response. The model held up remarkably well to a more detailed examination of limits reported in a series of papers by ROBERTS (1966*a*, 1966*b*, 1967*a*, 1967*b*). One restriction on the application of ROBERTS' results is that this work, like the early experiments, was also restricted to body weight.

The first paper in the series examines the limits ultimately reached in four large and three small lines, all selected for six-week body weight or else for growth between three and six weeks. The two characters are virtually identical in terms of the ranking of mice on the two measurements, as pointed out by FALCONER (1955). The earlier progress of these lines had been reported in various papers: KING (1950), FALCONER and KING (1953), FALCONER (1953, 1960*a*). The limits

were examined very largely in terms of ROBERTSON'S (1960) theoretical treatment of the topic. This treatment introduces the concept of the half-life of the selection response, as a measure of the time scale. It is of course quite impossible to determine the precise point at which an asymptotic curve reaches its maximum, but it is quite possible to estimate the point on the time scale by which half of the final gain has been achieved. The utility of the concept is that the half-life of a response is a function only of the effective size of the population, though the exact function will depend on whether all the genetic variance is additive (in which case the half-life is $1.4N$ generations, where N is the effective number); dominance and epistasis will lengthen the half-life. The half-life of the selection experiments analysed by Roberts were all of the order of $0.5N$ generations. As a rough arbitrary rule, we may therefore suggest that the duration of the response, when selecting for body weight in the mouse, lasts for a number of generations approximately equal to the effective size of the population. It is of interest to note that COMSTOCK'S results, cited earlier, fall into the same pattern. Although this rule is offerred only in a most tentative way, it may in fact be no accident. Firstly, the proportion of animals selected, in laboratory experiments with the mouse, is a quarter or a third of all animals measured. This is purely a function of reproductive rate in the mouse, if we operate on first litters only, but it happens to be a proportion close to the optimal in terms of maximising the final gain. Secondly, body weight in the mouse is a largely additive genetic trait, and is therefore free of some of the complications that arise in other situations. In any event, a short half-life means that there is a much reduced probability of the chance fixation of genes deleterious to the direction of selection, and this was one of the main conclusions from ROBERTS' (1966*a*) analysis. The other main conclusion was that the exhaustion of the additive genetic variance was a sufficient cause for the limits observed, although the analysis by no means excluded other causes.

Some empirical observations from ROBERTS' survey may also be mentioned briefly. The first is that different experiments with different stocks at different times all led to similar limits, measured on six-week weight. The limits appeared to be about 30 g. for large mice and 12 g. for the small ones, and the very best line means observed in each direction surpassed these figures only by a gramme or two. Secondly, about 20 *loci* each with an effect of around 0.75 phenotypic standard deviations would explain the selection results. Such estimates, by their nature, are arithmetically imprecise, and it may be intuitively more plausible to double or quadruple the number of genes and reduce their effects proportionately. But the trouble with trusting our intuition in that direction is that it makes it correspondingly more difficult to explain the very short half-lives of the responses, and these are incontrovertible experimental observations. Given those half-lives, we must accept genes with substantial efects.

In the second paper, ROBERTS (1966*b*) examined in more detail the genetic nature of the limit in two of the lines — one large and one small — included in the earlier analysis. The conclusion that the limit was due to the exhaustion of the additive variance was sustained in the case of the large line. Its mean did not alter even after eleven generations of reversed selection, proving conclusively the total absence of additive variance and presumably indicating that after 38 generations of selection for high body weight, fixation of genes affecting weight was

complete. In the case of the small line, however, the earlier conclusion had to be modified. Although it failed to respond to continued selection for small size, it gave some indication of a size increase when selection was relaxed, and more dramatically, it responded sharply to reversed selection, until its weight asymptoted again at about 18 g. The fact that the level of the base population was not regained proves that many of the original *loci* had been fixed for the allele giving small size, but there was plenty of residual segregation left to yield an astonishingly high value of 56 % for the heritability when the selection was reversed. The limit to downwards selection, in the absence of other detectable factors, was attributed to natural selection operating on viability, for which there was indirect evidence. Although it was not mentioned in the original paper, this may also be the explanation of the very high heritability when selection was reversed *i. e.* that the artificial selection for large size was being aided by natural selection operating in the same direction.

In the next paper, ROBERTS (1967*a*) re-examined the procedure successfully deployed previously by FALCONER and KING, described earlier. It was taken as axiomatic that lines from different (though overlapping) base populations would be fixed for different alleles at their limits, and that crosses between them ought to provide new genetic variance. Four large lines were available for experimentation, and though two of these derived from the same base, they had been selected on different diets, which might therefore have picked out different genes, as noted at the outset. The question was purely empirical: by combining the four lines and selecting from the 4-way cross, how much further advance could be obtained? The answer was, 25 % above the level of the largest of the original lines. All had reached their limits for six-week weight at 28 to 32 g, and after about 13 generations of selection from the cross, a new limit of 40 g was reached. It must be emphasised, when thinking of applications to animal production, that this result is not an argument for subdividing populations; that is another topic. But if, fortuitously or otherwise, distinct populations have reached their limits, the prognosis is good that further progress may be made by crossing them and continuing to select. A note of caution, however, was struck by the mouse experiment. Although the final gain was gratifying, there was no evidence that any gain had been made over the first six generations; if the same were to apply to species of domestic livestock, this would be a discouraging prospect. To explain the lag in response, ROBERTS invoked repulsion linkage, as a result of crossing, of genes affecting the trait. Linkage, as was shown by MCPHEE and ROBERTSON (1970), can have a considerable depressing effect on the total gain. Their experiment, however, was conducted with *Drosophila*, and linkage problems would not necessarily apply with the same force in the mouse, with its twenty chromosomes. ROBERTS' argument therefore demands some explanation. It was shown by HILL and ROBERTSON (1966) that the chance fixation of an unfavourable allele would be enhanced if it was linked to another *locus* affecting the trait; if the *loci* were of roughly equal effect, the chance of fixation was increased at both *loci*. *Loci* that are unlinked, on the other hand, would be mostly fixed for the favourable alleles. Now, turning the argument around, this would suggest that the *loci* where an unfavourable allele had been fixed would be predominantly those linked to other *loci* affecting the trait. It follows that in the crosses, the new genetic variance would derive from segregation among such «linkage groups», and only as the linkage broke

up would we obtain the desirable alleles on the same chromosome, to allow further advance. There was some evidence that this kind of phenomenon had occurred. A sample from the four-way cross had been random mated for seven generations before selection was applied to it. In this case, the response was immediate, unlike the initial selection directly from the four-way cross. Repulsion linkage in an organism like the mouse may therefore be a problem only in crosses between strains selected to their limits.

ROBERTS reported also on selection for decreased weight from a crossbred population derived from three small lines, which had also reached their limits. Small size, or reduced growth, is obviously of less applied interest, except where maintenance costs can be reduced without a deleterious effect on the marketable product. An example where such a system operates is provided by egg-laying strains of poultry. The mouse experiment, unlike the selection for increased weight, gave a largely negative result. After 24 generations of selection for a low weight from the crossbred, the mean had declined only marginally, and the low weights of two of the three lines which went into the cross had not been recovered. Obviously, by any applied standard, the procedure was an outright failure, and reasons for this must be sought. It is difficult to explain why the weights of the lowest line were not recovered; this is the minimal expectation, for all that had to be done was to reconstitute the genotype of that line from a population where its genes were at a frequency of at least 0.33, except for any *loci* that had not been fixed and which were not contributing to the response anyway. Dominance does not provide an adequate explanation. Directional dominance favours large size in the mouse, and selection for the recessive small ones should therefore have been efficient. Unlike the case of the large cross, linkage is not acceptable as an explanation either, unless it were exceptionally and unbelievably tight. By default of any other satisfactory hypothesis, we are left with the breakdown of epistatic combinations which, for some reason, could not be reconstituted. Even this explanation is not satisfying, for it leaves open the question of how those epistatic combinations of genes were assembled in the first place. Whatever the true explanation may have been, the formation of the crossbred population of small lines provides a cogent empirical case against the formation of gene pools to preserve genetic material. The blunt truth is that it may prove impossible to recover from the pool anything as good as what went into it.

Leaving aside this complication encountered with the small lines, the situation found in the large lines is clearly one of the exhaustion of the additive genetic variance, and further advance becomes possible only if a new source of genes can be tapped. In the final paper of the series, ROBERTS (1967*b*) reports two other experiments designed to introduce new genetic variance into a large line which had reached its limit for body weight. The first method employed X-irradiation, in an attempt to produce favourable mutants. The outcome was negative, for if there had been any gain, and that was doubtful, it certainly could not be considered worthwhile. It was concluded that the doses of irradiation required to provide reasonable hopes of success were too high to be tolerated by mammals. The result could not match the successful deployment of irradiation to gain further advance in *Drosophila*, reported by R. E. SCOSSIROLI (1954) and by R. E. SCOSSIROLI and S. SCOSSIROLI (1959). It was more in line with the experiences of CLAYTON and

ROBERTSON (1955, 1964), also with *Drosophila*, and of ABPLANALP, LOWRY, LERNER and DEMPSTER (1964) with chickens. The overall conclusion must be that irradiation has little to offer as a means of breaking through any limit in farm animals.

The second method of introducing new genetic variance into a line at its limit was, like LEWIS and WARWICK'S (1953) work, mentioned earlier, to cross it to an «unimproved» random-bred strain. For ease of interpretation, it would have been neater to outcross to the base population from which the selected line had been derived, but the base population by then did not exist. But the random-bred strain actually employed did overlap considerably with the extinct base population. The result was a modest success, at least in experimental terms. An advance of about 12.5 % over the original limit was obtained in two replicate lines to which the procedure was applied. Though the ultimate gain was substantial, by any standard, it was less successful than the other selection programme, from the crosses between the four lines at their limit. The comparison suffers further if we were to consider applying it to livestock. Firstly, there was again a lag of six generations before any improvement could be noted, and throughout this period, the performance of the outcross was well below that of the plateaued line. The level of the plateaued line was not recovered until the ninth generation, and the full gain was not recorded until generation thirteen or fourteen. All this suggests that before an analogous method could be even contemplated with farm animals, some careful costing would be necessary to evaluate the loss in current production against the hope of ultimate gain. Most definitely, «scrub» stock should not be considered if two or more breeds of roughly equal merit are available for crossing, for such a cross would avoid the initial depression in performance. The method may find some use if breeding and production stocks are distinct, and if the breeder has long-term aims. But in species like cattle, where the generation interval is long and where breeding stock must also pay for their keep, the prospects of using unimproved stock are not encouraging.

There has been a difficulty about the interpretation of my papers which I should now remove. In the second paper (ROBERTS, 1966*b*), I reported a sharp and unexplained rise in the mean of a large line at its limit, from a level of 32 g or so to around 35 g. In the paper, I argued against an environmental shift, and postulated a genetic change, either a new mutation or, more probably, a rare recombinational event. Ultimately, these two postulates become indistinguishable. But it is important to decide whether the change was environmental or genetic, as the evaluation of the other procedures depends critically on this decision. Subsequent events proved unambigously that the change was a genetic one, as supposed initially, and the evidence is presented briefly here. Dr. L. R. PIPER, now of C.S.I.R.O. EPPING, N.S.W., Australia, became actively involved in the resolution of this question, and I am much indebted to him for recording some of the data, and particularly for his critical appraisal of the problem.

The main experimental findings are summarised in Figure 1. The large line at its limit is the one designated *CL*, and until generation 43, it had been running at a fairly constant level of 32 g for over 20 generations. It then rose smartly to 35 or 36 g and remained at this higher level until generation 66, the last point shown in Figure 1. Just prior to the rise in weight, at generation 38, a branch had been taken from the *CL* line and kept without selection; this stock was designated *CLR*. Over 17 generations of random mabing, *CRL* kept close to the

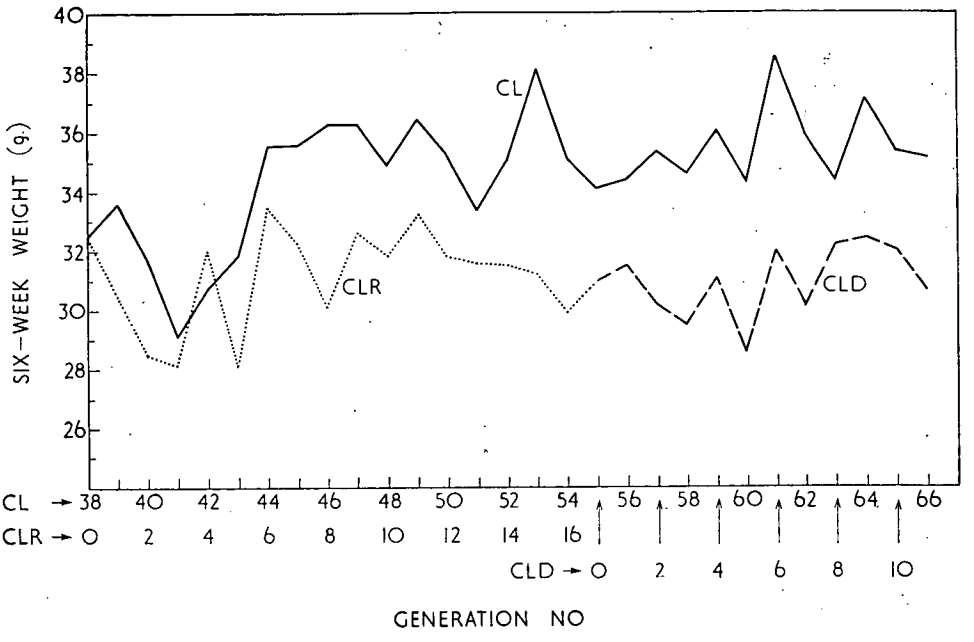


FIG. 1.

old limit of 32 g. Selection for largesize was then renewed, and the stock was redesignated *CLD*. Over 12 generations of further selection, no response was obtained. This proves conclusively (i) that *CL* had become genetically differentiated from *CLR/CLD* (ii) the genetic change which became apparent in *CL* at generation 44 was unique to that line.

Much work was done on these two lines, mostly by Dr. PIPER. Crosses between them, and backcrosses to the parental lines, always assumed a value intermediate between those of the immediate parents. The system behaved in a completely additive manner, as expected from a simple Mendelian model with no dominance. Attempts to demonstrate segregation, on the other hand, failed completely. This part of the programme was much too ambitious, and the reasons are discussed in detail by PIPER (1971). There is little hope of demonstrating the segregation of this kind of gene in any mammal, even if its effect is about one phenotypic standard deviation, as was the case here. Unfortunately, we did not test whether the *CL* line could be brought down again by reversed selection. Experience with the crosses makes it virtually certain that no response would have been obtained. The *CL* line behaved regularly as if whatever genetic change that had occurred had been fixed. Estimates of heritability in both lines were consistently compatible with zero.

This work did not resolve the question whether the genetic change had been a mutation or a recombinational event; the distinction is in any event a conceptual one. But that a genetic change had occurred in the *CL* line is no longer in

doubt. This facilitates the interpretation of the other experiments which hinged on this programme. It also suggests one possible way of breaking through a limit, namely to wait until something happens, but as a piece of practical advice, its futility is self-evident.

The experience from the work reviewed so far indicates that, although problems may arise, there is considerable hope of making further advance from an initial limit. Basically, what has been shown is that the limit is governed by the gene content of the material, which in turn is determined by the gene content of the base population and by chance fixation of genes during selection. To break through a limit, what we need therefore is a new source of genes, containing at least some alleles better than the ones we have already, to provide new genetic variance. Once that has been achieved, the other ingredients are non-genetic, namely, time and money. And it follows that selection limits in farm animals may well prove to be non-genetic limits, set by the high cost of long-term improvement programmes. Of the procedures discussed up to now, the most promising one, taking the laboratory mouse as a model, is the crossing of lines that have already been selected as far as they will go. In farm animal terms, these would be high-performance breeds or strains that have ceased to respond to selection for further gains. But if there is only one such breed available — and we can perhaps just about visualise an instance of that situation developing in the case of dairy cattle — where do we go? It has already been noted that crossing to inferior strains, in the hope of an ultimate gain many generations (and more years) hence, is not a practical proposition for some species. It is in this context that a paper by FALCONER (1971) is of particular interest, as it concerns specifically to the «one-strain problem»; it reports a method of improving that strain without stepping outside the boundaries of its gene content.

The experimental material was a line of mice selected to its limit for increased litter size. Though the selection had produced a line with 9.7 live-born offspring in first litters, the response had not been as impressive as had been the case in selection experiments for body weight; in the litter size experiment, the response had been only 1.8 phenotypic standard deviations, or 3.8 additive genetic ones. From the non-additive genetic nature of fertility, FALCONER argued that there could be a considerable amount of non-additive variance left in the line at its limit. He went on to test this hypothesis, and further, to explore whether such variance could be exploited to secure further gains.

Nine inbred lines of independent origin were set up from the strain, in its 42nd generation, each line being initially represented by about four sib pairs. During subsequent generations of inbreeding, 80 sib pairs were selected from the best litters (about 20) from the preceding generation. Selection occurred both between and within lines until the 7th generation, by which time the original 9 lines had been reduced to 4. Thereafter, 20 sib pairs were set up from best litters within each line, to ensure their survival. After 11 generations of sib mating, by which time the inbreeding coefficient had risen to 89 %, the four inbred lines were crossed according to a diallel scheme. The following generation was a 4-way cross, where all progeny had the four inbred lines equally represented in their parentage. These progeny formed a new line, which was continued over a further ten generations of random mating. Throughout these ten generations, the new strain showed a mean improvement in litter size of 1.54 ± 0.19 young per litter over the original

line at its limit, which was still being maintained. Thus, the method of selection during inbreeding proved to be very effective in improving litter size in a line which had reached its limit for the trait under normal selection.

To explain his results, FALCONER suggested that the improvement had been obtained through the removal of recessive genes, exposed by the inbreeding and removed by the selection, that had limited the performance of the line at its previous level. Theoretical considerations showed that the improvement could have been obtained by the removal of 30 such genes, each with an effect (homozygote difference) of 0.5 phenotypic standard deviations, and at frequencies of 0.2 in the line at its original limit. It was further calculated that the procedure would have removed 75 % of the segregating recessives, which therefore puts their number at 40 when the line first plateaued. The additive variance generated by such genes is still compatible with estimates of zero for the realized heritability.

An incidental but important observation noted by FALCONER was that his best lines depressed very little on inbreeding, while the best survived 20 generations of sib-mating at a level fully equal to that of the line at its limit. This rules out overdominance as an important cause of the residual segregation. Should overdominance prove to be a problem in practice, the reader is referred to the wide-ranging discussion provided by BELL, MOORE, and WARREN (1955) of the relevant breeding techniques.

The generality of the efficacy of FALCONER's method of selecting while inbreeding was tested further by AL-MURRANI and ROBERTS (in press) on a line of mice that had reached its limit for high body weight. Fertility problems meant that they had to stop inbreeding at 50 %, but otherwise the procedure was similar. Partly perhaps because the inbreeding stopped so soon, but more probably because of the difference in the genetic nature of the trait, the method failed to increase body weight. Nevertheless, significant differentiation between inbred lines drawn from the line at its limit, proved that the fixation of genes affecting body weight had not been complete. This contrasts with a line with a longer history of selection for high body weight (ROBERTS, 1966*b*), noted earlier. As in FALCONER's case with litter size, the segregation of recessives at the limit could also explain AL-MURRANI and ROBERTS' results with body weight. Their system could tolerate perhaps as many as 10 *loci*, with effects of about 0.67 phenotypic standard deviations (or some equivalent combination of numbers and effects) and, again, at frequencies around 0.2. Nevertheless, in this case, the total improvement to be gained by their complete elimination was only about 2 % over the level of the original limit. The observation is not new, but AL-MURRANI and ROBERTS' results illustrate how substantial genetic effects can occur at individual *loci* despite trivially low heritabilities and negligible potential gains. This prompts the suggestion that any endeavours to eliminate recessives from populations of farm animals should be considered very carefully; it is quite possible that the effort may not be worthwhile.

There is no reason to suppose that long-continued selection would not ultimately remove segregating recessives from a population, as indeed it should. If selection does not drive them to fixation, drift no doubt will. But, as is widely appreciated, the selection becomes increasingly ineffective as the frequencies of the recessives fall. In characters like litter size, and other measures of fertility, which have large non-additive components, the residual recessives may seriously

depress the final limit. In such cases, FALCONER's scheme of selecting with inbreeding may well find an application. CLAYTON (1972*a*) commends the scheme to the attention of poultry breeders, who may already be in a situation where it could perhaps be employed. In the more additive traits, like body weight, the scheme has obviously less appeal, and with the slower breeding livestock, where each newborn young has a high cash value, inbreeding could never be seriously advocated.

One final point has to be drawn from the work on selection limits in the laboratory mouse. Most of this work, as is by now obvious, has been conducted on body size. Inevitably, it seems, selection for either increased or decreased growth has led to fertility problems. Small mice have small litter sizes; large mice certainly have larger litters, but unfortunately, fewer of them. Sterility is common place among large mice (see, for instance, BRADFORD, 1971) and the length of their reproductive life is much reduced (ROBERTS, 1961). FOWLER and EDWARDS (1960) report on some of the factors reducing fertility in both large and small mice, and their experiences are unhappily only too widely shared. Part of the trouble may be the small size of laboratory populations, leading to accumulated inbreeding. But this reason alone is insufficient; FALCONER (1960) reports on a random-bred stock where the accumulated inbreeding over 31 generations was 32 %, and where there had been no detectable effects on litter size. It could be argued that selection for any trait, involving as it does changes in gene frequencies and the tendency to fix chromosomal segments, might be expected to have a deleterious effect on fitness. But over and above any general difficulties associated with selected lines, there seem to be particular problems among those selected for size — large or small. To the extent that we may generalize from mouse experiments, fertility problems may well arise in farm animals long before we need become concerned with the genetic nature of the limit to artificial selection. CLAYTON (1972*b*) reports on the reduced fertility of several avian species selected for more rapid growth, indicating that the problem is not confined to mammals.

Even among laboratory mice, infertility may well be the immediate cause of limits. FALCONER (1974) reports on six small lines which were responding only very slowly, in absolute terms, because he was running out of selection differentials, as a result of a poor reproductive rate. ROBERTS (1967*a*) lost two of his largest lines through infertility. In his case, females were already becoming too fat to breed when they were mated at 6 weeks of age, or as soon as possible thereafter, according to normal mouse practice. An offshoot of one of these large lines was taken just before the line became extinct, and this offshoot was mated a week earlier, at 5 weeks. At this age, the accumulation of fat was not sufficient to cause infertility, and because of their large size, the animals were also sexually mature at an earlier age. The line mated at 5 weeks survived without trouble for a further 23 generations, and its extinction was artificial and deliberate. In its terminal generation, 14 of the 15 pairs set up proved fertile, with a mean of 9.7 young born per first litter. This kind of performance tends to rule out the accumulated effects of inbreeding in small populations, and implicates other factors as the cause of fertility problems in selected stocks.

To the extent that fertility may prove to be a problem in improved breeds of domestic livestock, the solution may therefore not be too difficult. Early mating is probably adopted routinely where selection for rapid growth results in earlier

sexual maturity. It is also easier to restrict food intake in domestic livestock than it is with an animal as small as the mouse. This may completely overcome fertility problems associated with fatness. Again, food intake regulation of breeding stock is already the established practice in the more advanced systems of animal production. But there may still be a residue of breeding difficulties with some improved strains. For this reason, if for no other, the commercial animal in such cases will very likely be the product of some crossing system.

What has been attempted in this review is to examine the causes of selection limits in laboratory mice, and to explore methods of further advance, with some attention to the relevance of such methods in the context of animal breeding. It is perhaps encouraging to note that substantial further gains have been reported for growth and fertility, as the corresponding traits in farm animals are of obvious and direct economic importance. What is discouraging is the time required to secure such gains, and the potentially high cost of some of the breeding methods employed if they were applied to domestic levestock. In view of this, breeding objectives should perhaps be redefined more clearly in terms of the efficient exploitation of available nutritional resources. It may be none too soon, or not much too soon, to de-emphasize selection for growth or fertility *per se*, if gains in these traits make increasing demands on the nutrients available for direct consumption by man. The animal of the future may be the efficient converter of the foods which man can not, or will not, eat; particularly, the animal may have to be an efficient harvester of crops in areas which, for a variety of reasons, would otherwise go to waste. Simultaneously, the advent of meat subtitutes redirect attention towards the quality and palatability of the marketable product, which in some species has clearly suffered from selection for rapid growth and, correspondingly, younger slaughter ages. It is not too difficult to imagine that some modern broiler chickens, despite their low price, may be vulnerable if an even cheaper meat-like spun protein competes with them as a bland base for various sauces. In other ways too, the marketable products may change. The energetic cost of transporting liquid milk, for instance, may present the breeders of dairy cattle with new objectives.

None of these arguments arise directly from a discussion of selection limits, but experiences with breaking limits in the laboratory mouse certainly reinforce them. They show clearly that conventional breeding systems lead to a point where further progress, though perfectly possible, may become uneconomic. It is suggested that breeding aims can profitably be pushed only so far. As the aims become redefined, whether in response to genetic or economic needs, no doubt those aims too will ultimately be driven towards their limits. The laboratory mouse has an obvious role to play in exploring some of these new systems, and work of the kind reported by SUTHERLAND *et al.* (1970), on selection for appetite and efficiency, will no doubt be extended. In this way, we may gain better understanding of how animals may be adapted to meet changing needs, and at least some qualitative assessment of the potential advance.

SUMMARY

Experimental evidence from laboratory mice suggests that the limit to selection for high growth rate will be determined largely by the fixation of genes affecting

the trait. Any residual segregation at unfixed *loci* will have trivially small effects on the trait. To achieve further gains, new genetic variance must therefore be introduced from some other source. The most sucessful method reported so far is to select further from crosses between lines previously selected to their limits.

Selection for high fertility reaches its limit while a large number of deleterious recessive genes still segregate. Considerable further advance is posible if these genes are eliminated. This may be achieved by inbreeding, selecting the best inbred lines, and then crossing those lines.

Although further gains are thus quite feasible under laboratory conditions, for both growth and fertility, the applications of the methods to domestic livestock would demand heavy investment, both in time and money. It is suggested that it may therefore be fortunate that breeding objectives are likely to change, in most species of livestock, before current schemes for improving growth and fertility attain their limits.

RESUME

Les résultats obtenus avec les expériments de sélection ayant comme but l'augmentation de l'index de croissance suggérent que les limites de la sélection sont dûs à la fixation des gènes qui controllent le caractère sélectionné. Après avoir arrivé à un limite il est possible encore certaine segrégation résiduelle, dont son efect sur le caractére sélectionné manque de significance. Donc, le déferlement de ces limites sera seulement possible à travers de l'introduction de nouveau matérial génétique chez la population sélectionnée à limite. Jusqu'au cette date, le croisement entre des lignées sélectionnés au limite, suivie d'une nouvelle selection dans le même croisement, a été la méthode qui a donné des meilleurs résultats en ce qui concerne la déferlement de ce type de limites à la sélection.

Au contraire, quand on a arrivé au limite chez des lignées sélectionnés pour un haut niveau de fertilité, un considerable nombre de gènes récessifs d'effet nocive sur le caractère sélectionné continuent leur segregation. No obstant, on peût arriver à un limité plus élevé à travers de l'élimination de ces gènes, en établissant des lignées consanguines à partir de la population sélectionnée dans le limite et de la sélection des lignées les meilleures, suivie par des croisements entre elles.

Malgré la déferlement des limites à la sélection pour l'index de croissance et fertilité est parfaitement possible dans les expériments du laboratoire, l'application de cettes méthodes chez des populations domestiques aurait besoin de grands investissements économique a long terme. Etant donné qu'il est probable que les objectifs de l'améliorement de la plus grande partie des espèces domestiques changeront avant que les systèmes actuels d'améliorement arrivent aux limites à la sélection, on mentionne dans le travail que, etant données les difficultés de déferlement des limites exposes, ces changes pourront éviter la mis au point des dites méthodes.

RESUMEN

Los resultados obtenidos en experimentos de selección para elevar el índice de crecimiento sugieren que los límites a la selección se deben a la fijación de los genes que controlan el carácter seleccionado. Una vez alcanzado un límite, aún

puede observarse cierta segregación residual, cuyo efecto sobre el carácter seleccionado es insignificante. Por tanto, la rotura de estos límites sólo será posible mediante introducción de material genético nuevo en la población seleccionada al límite. Hasta la fecha, el cruzamiento entre líneas seleccionadas en el límite, seguido de nueva selección en dicho cruzamiento, ha sido el método que ha proporcionado mejores resultados en lo que respecta a la rotura de este tipo de límites a la selección.

Por el contrario, al alcanzarse el límite en líneas seleccionadas para alto nivel de fertilidad, un número considerable de genes recesivos de efecto perjudicial sobre el carácter seleccionado continúan segregando. Es, sin embargo, posible alcanzar un límite más elevado mediante eliminación de estos genes, lo que puede conseguirse estableciendo líneas consanguíneas obtenidas a partir de la población seleccionada en el límite y selección de las mejores líneas seguida por cruzamientos entre éstas.

Aunque la rotura de los límites a la selección para los índices de crecimiento y fertilidad es perfectamente posible en experiencias de laboratorio, la aplicación de estos métodos en poblaciones domésticas requeriría grandes inversiones económicas a largo plazo. Como es probable que los objetivos de la mejora de la mayor parte de las especies domésticas cambien antes de que los esquemas actuales de mejora alcancen los límites a la selección, se indica que, dadas las dificultades de rotura de límites expuestas, estos cambios pudieran evitar la puesta en práctica de dichos métodos.

REFERENCES

ABPLANALP, H.; LOWRY, D. C.; LERNER, I. M., and DEMPSTER, E. R. (1964): Selection for egg number with X-ray-induced variation. *Genetics*, 50:1083-1100.

AL-MURRANI, W. K., and ROBERTS, R. C. (1974): *Genetic variation in a line of mice seleced to its limit for high body weigth.* (In press.)

BATEMAN, N. (1971): Effects of diets consisting of maize and milk on growth and carcass characteristics of mice. *Anim. Prod.*, 13:413-424.

BRADFORD, G. E. (1971): Growth and reproduction in mice selected for rapid body weight gain. *Genetics*, 69:499-512.

BUTLER, L. (1952): A study of size inheritance in the mouse. II. Analysis of five preliminary crosses. *Canad. J. Zool.*, 30:154-171.

CLAYTON, G. A. (1972a): Selection plateaux in poultry. *Ann. Génét. Sél. Anim.*, 4:561-568.

CLAYTON, G. A. (1972b): Effects of selection on reproduction in avian species. *J. Reprod. Fertil.*, Suppl., 15:1-21.

CLAYTON, G. A., and ROBERTSON, A. (1955): Mutation and quantitative variation. *Am. Nat.*, 89:151-158.

CLAYTON, G. A., and ROBERTSON, A. (1964): The effects of X-rays on quantitative characters. *Genet. Res.*, 5:410-422.

COMSTOCK, R. (1973): Growth in mice. *Genetics*, Suppl., 74:s51-s52 (abstract). Proc. XIIIth Int. Congr. Genet. (Berkeley, 1973).

DICKERSON, G. E. (1955): *Genetic slippage in response to selection for multiple objetives.* Cold. Spr. Harb. Symp. Quant. Biol., 20:213-224.

DICKERSON, G. E. (1963): Experimental evaluation of selection theory in poultry. In *Genetics Today* (Vol. 3, pp. 747-761). Proc. XIth Int. Congr. Genet. (The Hague, 1963).

FALCONER, D. S. (1952): The problem of environment and selection. *Amer. Nat.*, 86:293-298.

FALCONER, D. S. (1953): Selection for large and small size in mice. *J. Genet.*, 51:470-501.

FALCONER, D. S. (1955): *Patterns of response in selection experiments with mice.* Cold. Harb. Symp. Quant. Biol., 20:178-196.

FALCONER, D. S. (1960a): Selection of mice for growth on high and low planes of nutrition. *Genet. Res.*, 1:91-113.

FALCONER, D. S. (1960b): The genetics of litter size in mice. *J. Cell Comp. Physiol.*, 56:153-167.
FALCONER, D. S. (1971): Improvement of litter size in a strain of mice at a selection limit. *Genet. Res.*, 17:215-235.
FALCONER, D. S. (1974): Replicated selection for body weigth in mice. *Genet. Res.* (In press.)
FALCONER, D. S., and KING, J. W. B. (1953): A study of selection limits in the mouse. *J. Genet.*, 51:561-581.
GOODALE, H. D. (1938): A study of the inheritance of body weight in the albino mouse by selection. *J. Hered.*, 29:101-112.
GOODALE, H. D. (1941): Progress report on possibilities in progeny-test breeding. *Science*, 94:442-443.
HILL, W. G., and ROBERTSON, A. (1966): The effect of linkage on limits to artificial selection. *Genet. Res.*, 8:269-294.
KIMURA, M. (1957): Some problems of stochastic processes in genetics. *Ann. Math. Statist.*, 28:882-901.
KING, J. W. B. (1950): Observations on the mutant «pygmy» in the house mouse. *J. Genet.*, 53:487-497.
LERNER, I. M. (1950): *Population Genetics and Animal Improvement*, pp. 342. Cambridge, Cambridge University Press.
LEWIS, W. L., and WARWICK, E. J. (1953): Effectiveness of selection for body weight in mice from inbred and outbred populations derived from common parent stocks. *J. Hered.*, 44:233-238.
LUSH, J. L. (1945): *Animal Breeding Plans*, 3rd Ed., 443 pp. Ames, Iowa: Collegiate Press.
MCPHEE, C. P., and ROBERTSON, A. (1970): The effect of suppressing crossing-over on the response to selection in *Drosophila melanogaster*. *Genet. Res.*, 16:1-16.
MATHER, K. (1949): *Biometrical Genetics*, 162 pp. London, Methven.
OSMAN, H. S., and ROBERTSON, A. (1968): The introduction of genetic material from inferior into superior strains. *Genet. Res.*, 12:221-236.
PIPER, L. R. (1971): *The isolation of genes underlying continuous variation*. Ph.D. Thesis, University of Edinburgh.
ROBERTS, R. C. (1961): The lifetime growth and reproduction of selected strains of mice. *Heredity*, 16:369-381.
ROBERTS, R. C. (1966a): The limit to artificial selection for body weight in the mouse. I. The limits attained in earlier experiments. *Genet. Res.*, 8:347-360.
ROBERTS, R. C. (1966b): The limit to artificial selection for body weigth in the mouse. II. The genetic nature of the limits. *Genet. Res.*, 8:361-375.
ROBERTS, R. C. (1967a): The limit to artificial selection for body weight in the mouse. III. Selection from crosses between previously selected lines. *Genet. Res.*, 9:73-85.
ROBERTS, R. C. (1967b): The limits to artificial selection for body weight in the mouse. IV. Sources of new genetic variance - irradiation and outcrossing. *Genet., Res.*, 9:87-98.
ROBERTSON, A. (1960): A theory of limits in artificial selection. *Proc. R. Soc. B.*, 153:234-249.
SCOSSIROLI, R. E. (1954): Effectiveness of artificial selection under irradiation of plateaued populations of *Drosophila melanogaster*. *Symp. Genetics of Population Structure* (Pavia). (A. A. Buzzati-Traverso, ed.), pp. 42-66.
SCOSSIROLI, R. E., and SCOSSIROLI, S. (1959): On the relative role of mutation and recombination in responses to selection for polygenic traits in irradiated populations of *Drosophila melanogaster*. *Int. J. Radiat. Biol.*, 1:61-69.
SUTHERLAND, T. M.; BIONDINI, P. E.; HAVERLAND, L. H.; PETTUS, D., and OWEN, W. B. (1970): Selection for rate of gain, appetite and efficiency of food utilization in mice. *J. Anim. Sci.*, 31:1049-1057.

PAPER 'O.

Genetic variance in a line of mice selected to its limit for high body weight.

Anim. Prod. 19, 273-289. 1974.

by

W.K. AL-MURRANI & R.C. ROBERTS

Anim Prod. 1974, 19: 273–289

GENETIC VARIATION IN A LINE OF MICE SELECTED TO ITS LIMIT FOR HIGH BODY WEIGHT

W. K. AL-MURRANI† and R. C. ROBERTS‡
Institute of Animal Genetics, West Mains Road, Edinburgh, EH9 3JN

SUMMARY

A line of mice, at its limit to selection for high body weight did not decline in performance over 11 generations of random mating, neither did it respond when selection was renewed. The experiment tested a method of improving body weight by a scheme which had earlier increased litter size under similar circumstances. The scheme was to derive partially inbred lines from the plateaued line, to select during inbreeding and, finally, to cross the best inbreds. Body weight was not increased, but the study allowed further examination of the residual genetic variance in the line.

During inbreeding, the inbred lines became clearly differentiated in body weight, proving that loci controlling body weight had not become fixed. There was also a significant response to selection for a lower body weight during inbreeding. The pattern of results suggested the segregation of recessive genes, detrimental to high body weight but which selection had become inefficient at removing. A genetic model compatible with the results accommodated several such recessives, perhaps as many as 10, each with an effect of about two-thirds of a standard deviation (or some equivalent combination of gene number and effect), and at frequencies of around 0·2. Nevertheless, the total improvement in body weight to be gained by their elimination was only half a gram, or less than 2%. Thus, substantial genetic effects can occur at individual loci despite trivially low heritabilities and negligible potential gains.

INTRODUCTION

THE commonest method of breaking through a limit to artificial selection is to introduce, from some source, new genetic variance to render the trait amenable to further selection. Such new variance is probably most easily obtained by crossing a plateaued population to another strain, and examples of this approach, with specific reference to body weight in the mouse, have been provided by Falconer and King (1953) and by Roberts (1967a and b). In contrast to this method, Falconer (1971) reported a considerable advance over an earlier limit by exploiting the residual genetic variance within a line. He improved litter size by 1·5 young per litter in a line of mice that had long ceased to respond to selection for that trait. His method was to derive a number of partially inbred lines from the plateaued population,

† Present address: College of Veterinary Medicine, Abu-Ghuraib, Baghdad, Iraq.
‡ Agricultural Research Council, Unit of Animal Genetics.

selecting on individual merit during inbreeding and finally, to cross the best of the surviving inbreds.

Falconer's method, despite the cost and tedium of inbreeding, suggests one important advantage over the alternative method of outcrossing. The outcrossing method requires lines of equal merit for crossing; otherwise, average performance will initially be less than the level of the better line until this level is restored, and hopefully exceeded, by selection. For additive traits, it might therefore be difficult to improve one superior breed or strain without initially sacrificing some performance. It is of interest to note that among Falconer's inbreds, despite the expected general depression found on inbreeding, the better of his inbred lines depressed very little, while the best line did not depress at all. Should it generally prove feasible to select for good performance among the better inbred lines, Falconer's method of breaking through a limit might be of applied interest, should one breed or strain be too far ahead of its competitors to tolerate the loss in performance on crossing.

The successful application of Falconer's method demands the presence of residual nonadditive genetic variance after the additive variance has been exhausted. Falconer's experiment was on litter size, which is a character known to have a large nonadditive component. One purpose of this paper is to explore the generality of the method, by applying it to a character where nonadditive genetic variance is of less relative importance. One such character is body weight, which has the additional relevance, like litter size in Falconer's experiment, of economic interest in domestic livestock. A line of mice, at its limit to selection for body weight, was available for experimentation. Apart from the possible improvement in body weight, the method would allow us to examine how much genetic variance remained in the line, and also perhaps to determine something of its nature.

MATERIAL AND METHODS

The general outline of the experiment was as follows. A line of mice, designated QLA, had been selected to its limit for body weight at 6 weeks of age. From this line, 20 sib pairs of mice were obtained, each sib pair being mated to form the basis of one inbred line. During further inbreeding, selection was continued for high body weight at 6 weeks, both within and between lines. After three full-sib matings, four of the best lines among those that survived the inbreeding were crossed, over two generations, to form four-line crosses. The essential comparison was between the four-line crosses and the oringinal population.

The formation and selection history of the QLA line is given by Falconer (1974). Each generation initially comprised eight matings. At generation 23, the number of matings was increased to 16 and selection was suspended. There was no further selection for body weight between generation 23 and 29, when the line provided mice for the present experiment. We shall refer to the base population at this point as QLA 29, and it is the ancestral stock of all the derived lines described below.

Since body weight at 6 weeks is significantly affected by the size of the litter in which that mouse was born and reared, each individual 6-week weight was corrected to a mean litter size of eight at weaning, when the mice were 3 weeks of age. The correction was applied from the regression

of the mean weights of individual litters (sexes averaged) at 6 weeks on litter size at weaning, calculated from all the litters from the preceding nine generations of the QLA line, giving a total of 89 litters. The regression was $-0{\cdot}856 \pm 0{\cdot}122$ g/mouse, and the correction factor adopted was $\pm 0{\cdot}86$ g for each unit deviation from the arbitrary standard litter size of eight. There was no evidence of any curvilinearity of the regression over the range of litter sizes for which there were adequate numbers (from 4 to 14), and neither was there any suggestion of heterogeneity in the regression coefficient between generations, which were therefore pooled.

Although the above correction factor was calculated for outbred mice, there was no option but to apply it to the inbred mice as well, when inbreeding began. Retrospectively, we were able to test its appropriateness, by calculating the same regression with inbred litters. There was no indication that the regression varied systematically with the degree of inbreeding, and the value of the pooled regression from the inbreds was $-0{\cdot}968 \pm 0{\cdot}152$. Given rather large standard errors, the agreement between the two regressions is satisfactory.

As an added precaution against overweighting the larger litters, all generation means were calculated as the mean of litter means, whereby each litter was weighted equally. The only complication was the sporadic case of litters comprising one sex only. In such cases, half of the mean difference in weight between the sexes, for that stock and that generation, was added or subtracted, as appropriate.

Five stocks of mice were developed, all deriving directly or indirectly from QLA 29. The purpose of each stock and its breeding will now be described.

(a) *Selection for increased 6-week weight* (LAU)

As the base population had not been selected in its recent history, selection was renewed in order to test whether any further improvement in 6-week weight could be obtained by this method alone.

Sixteen matings were set up in each generation and a within-family method of selection was applied, identical to that described by Falconer (1974) for the initial selection. The largest mice of each sex were selected from each family at 6 weeks of age, and mated at random, save only for the avoidance of close relatives. The LAU (selected) stock was kept in step with the QLA stock (previously selected, but now random mated) from which it derived.

(b) *Inbreeding without selection* (IC)

As the QLA base population was at or near its limit to selection for body weight, it was to be expected that many (perhaps most) of the genes affecting weight had been fixed. This implies homozygosity of many chromosomal segments, involving genes other than those affecting weight. The effect of inbreeding in such populations is not well documented, and the purpose of the IC stock was to measure the effects of inbreeding, without selection, on body weight. This information was needed to evaluate the main programme, described under (c) below.

Sixteen sib pairs were taken at random from the 31st generation of QLA,

each sib pair forming one inbred line, and continued by brother-sister mating. The 16 lines were derived from 12 different families in QLA 31.

(c) *Inbreeding and crossing with selection* (IL)

This stock constituted the main experiment, outlined at the beginning of this section. Twenty inbred lines were set from QLA 29. Twelve lines were founded by one sib pair from each family; the other eight sib pairs derived as a second pair from some of the larger families.

There followed three generations of full sib mating, while selection was applied both within and between lines. In the process, some lines became extinct, while others became represented by multiple sublines, in order to keep the number of parental pairs constant at 20. The progress and fate of various lines is summarized in Figure 1.

TABLE 1

The numerical contribution of designated inbred lines to the crosses

Line number	Mean corrected 6-week wt (g)	No. of families in line	Whether used for crossing	No. of mice contributed		
				♀	♂	Total
1	30·80	2	Yes	7	5	12
3	28·29	3	No			
5	30·69	5	Yes	9	3	12
6	29·87	2	Yes	4	8	12
8	31·70	1	No			
13	30·49	1	Yes	4	8	12

The method of selection was between sib pairs without regard to their origin. Each mouse was weighed at 6 weeks of age, its weight being adjusted for the size of the litter in which it was born, as was described earlier. Within each litter, the heaviest male was nominally paired with the heaviest female, and their mean adjusted weight calculated. Then the next heaviest pair were identified and so forth, until all available pairs had been listed. This procedure was applied to all first litters that had been born within a reasonable time. When all the information became available for a generation, the pairs were ranked without regard to origin, and the 20 heaviest were selected to continue the inbreeding. There was a nominal constraint that the number of lines of independent origin should not be less than four, but in practice, no manipulation of the system became necessary to ensure this.

After three generations of brother-sister mating, four of the remaining six lines, or rather groups of sublines, provided parents for the first cross. The mean weight of each line and the number of mice of each sex that it provided for crossing is shown in Table 1. Unfortunately, the heaviest line (shown in the table as line number 8) could not be utilized in the cross, because of inadequate numbers. The lightest line (line number 3) was also rejected, while the remaining four lines were crossed to form the six possible crosses. There were four matings per cross, and each cross had both reciprocals represented, though not necessarily equally because of difficulties with numbers. This first cross was designated IX 1. The subsequent cross,

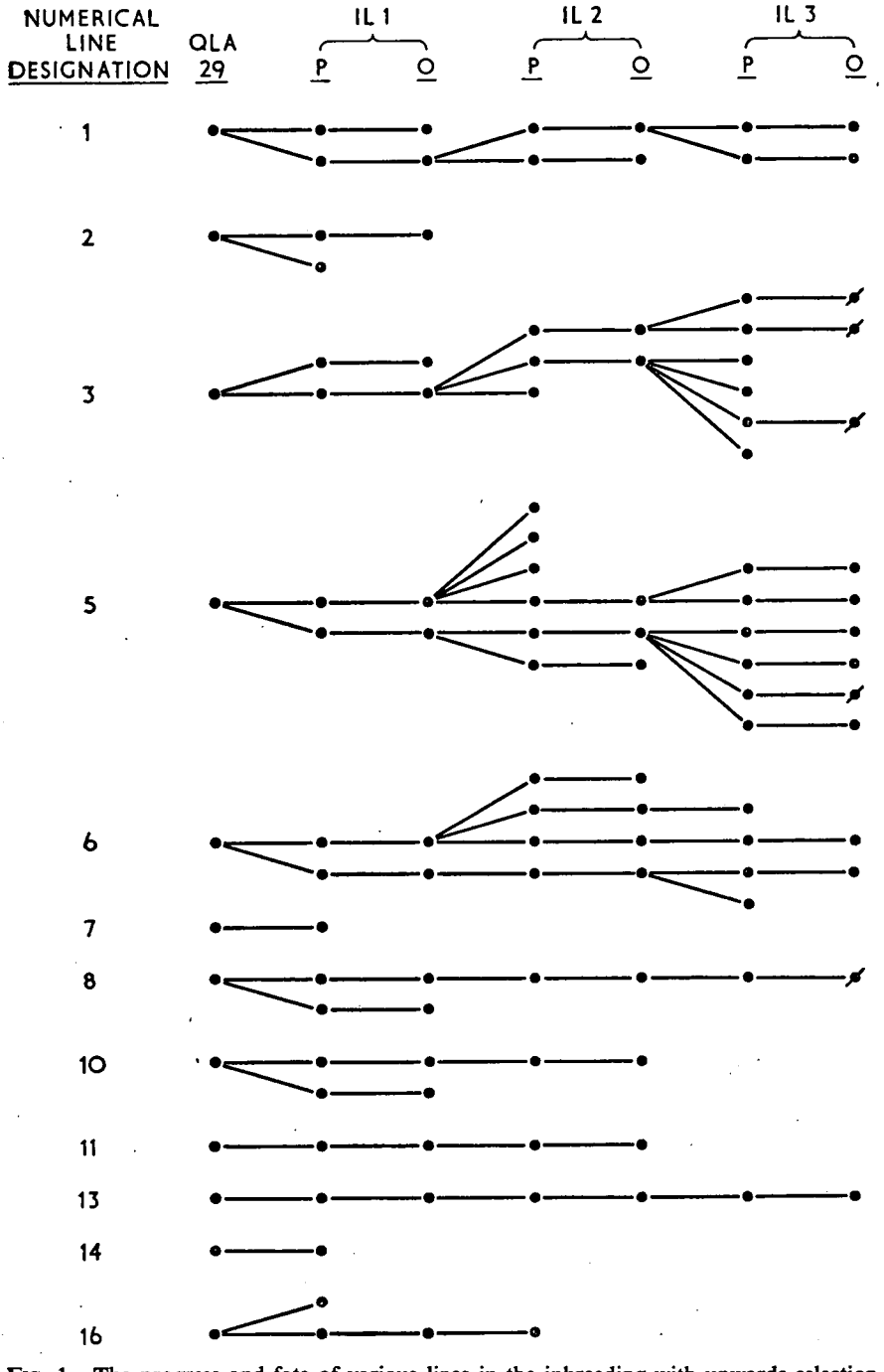


FIG. 1. The progress and fate of various lines in the inbreeding with upwards selection (IL) programme. P = parents and O = offspring. Sublines whose terminal points occur in columns headed P became extinct through sterility. The other extinct sublines either failed to produce offspring in time or else did not yield viable offspring of each sex. Sublines with an oblique bar through their terminal points in the last column did not contribute to the crosses.

IX 2, also comprised 24 matings, arranged to ensure that each offspring of this cross had the four parental lines equally represented in its ancestry.

The inbreeding of parents and of offspring in the separate generations are summarized thus:

	Inbreeding coefficient	
Generation	Parents	Offspring
QLA 29	0	0
IL 1	0	0·25
IL 2	0·25	0·375
IL 3	0·375	0·50
IX 1	0·50	0
IX 2	0	0

It would have been desirable to attain a higher coefficient of inbreeding than the 0·50 actually reached, but already there were difficulties in maintaining the lines. This was not entirely unexpected, for even without inbreeding, the productivity of this stock of mice was not very good.

(d) *Partial replication of the main experiment* (RIL)

By generation IL3, certain aspects of the data indicated the need for more information, and we decided to run a partial replicate of the main experiment. In retrospect, it is now clear that the replication should have been more complete, if it was to have full value. However, despite their deficiencies, the additional data obtained from the replicate augment the main conclusions, and they will be described briefly later.

Proper experimental design would have dictated that we should have returned to the base population (QLA), and commenced inbreeding another 20 lines. For reasons of convenience, we used mice from LAU, which by then had been separated from QLA by three generations of selection. This flaw, as will become apparent, was probably a trivial one. A more serious flaw was that we took only 10 new inbreds, and poor breeding performance reduced their number to three by the second generation of inbreeding. By the next generation, one of these had been reduced to one sex. In general, the data from the replicate, RIL, became increasingly unsatisfactory, and the second cross (corresponding to IX 2) was not even attempted.

(e) *Inbreeding with downward selection* (IS)

This line of mice represents another supplementary study which, if its value had been properly anticipated, might have been better designed. The main experiment (IL) was, as described earlier, limited to selection for higher body weight while inbreeding proceeded. The effect of the selection could clearly be more accurately assessed from a scheme of divergent selection. After one generation of upward selection, the lightest pair from eight of the original lines were also mated, and these were the founders of a downwards line, also selected under inbreeding. The scheme of ranking sib pairs was adopted, as in the upwards line (IL), except that in the case of IS, the smallest pairs were selected. The line continued for two further

generations, i.e. until the inbreeding coefficient reached 0·50. It was then discarded.

The main deficiency of the IS programme was its small size. It would also have been preferable to start the downwards selection by reverting to the outbred base population, rather than from the first upwards selected generation. But we considered it more important to keep mice of equivalent inbreeding as contemporaneous as possible. Any significant divergence under these conditions could be more clearly interpreted.

RESULTS

The main experiment, testing the effectiveness of selection in the presence of inbreeding, cannot be interpreted without the independent assessment of the two procedures. These will therefore be described first, and the results will at the same time indicate the extent of any residual genetic variance in the base population (QLA 29).

The independent effect of selection (LAU)

The results of renewed within-litter selection for 6-week weight (LAU) were compared over several generations with the original QLA stock, where mating continued to be at random. The data are summarized in Table 2.

TABLE 2

Mean corrected 6-week weights (grams) of the original line (QLA), and in the line where selection was renewed (LAU), and in the derived unselected inbreds (IC). Rows contain contemporaneous generations—see *text for origin of various lines*

Generation number	Relaxed QLA	Selected LAU	Generation number	Inbred IC
29	30·61 ± 0·85	—	—	—
30	30·96†	28·16 ± 0·88	—	—
31	29·11	29·42 ± 0·69	—	—
32	31·65	30·85 ± 0·68	1	28·82 ± 0·97
33	31·15	29·72 ± 0·85	2	30·89 ± 0·79
34	31·00	30·50 ± 0·45	3	29·71 ± 0·46
35	30·85	30·85 ± 0·98	4	29·16 ± 1·19
36	—	29·22 ± 0·69	—	—

† Standard errors for QLA 30 to 35 similar to those shown for LAU.

The standard errors for the generation means of QLA were roughly the same as those shown for LAU. In other words, a difference between the means of more than 2 g would be necessary if it were to approach statistical significance. None of the differences attains this magnitude, and furthermore, there is no hint of a divergence between the selected and random-mated populations. Indeed, if anything, the selected line tends to be lower than the relaxed line.

We can add to the information some of the previous history of the QLA line, and the means from generation 23 to the time corresponding to the end of the present study are shown in Figure 2. Although the mean weights fluctuate in a manner characteristic of small samples, there is no consistent or significant trend with time. If we ignore some high means at generations

24 and 25, the corrected body weights are remarkably steady at around 32 g. From this, we must conclude either that natural selection is inoperative or else that there is no additive variation in fitness, as it affects body weight remaining in the line. The failure of renewed selection (LAU) to increase body weight, noted earlier, points clearly to the absence of any detectable additive variance in body weight itself.

In support of this last conclusion, the following are the results of some calculations based on the data from LAU:

(i) The realized heritability, estimated from the regression of generation means on cumulated selection differentials, was 0·017±0·132. The

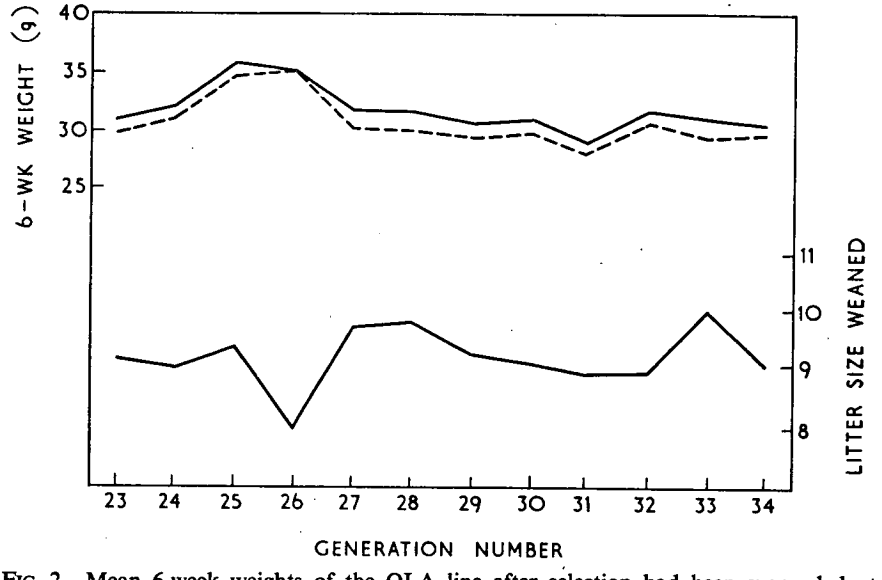


FIG. 2. Mean 6-week weights of the QLA line after selection had been suspended at generation 23. Solid line—weights corrected for litter size effects; broken line—weights not corrected. The corresponding litter sizes at weaning are shown in the lower graph.

estimate is totally insignificant, even though the standard error is biased downwards (Hill, 1972).

(ii) Likewise, the regression of body weights of offspring on those of their sires (dams being omitted to avoid maternal effects) was 0·014± 0·121, when pooled over the six generations. This regression of course estimates half the heritability, and is obviously compatible with the realized heritability, and with zero.

(iii) The failure of response to selection was not due to the lack of a selection differential. Over the six generations, the realized selection differential, when cumulated, was 9·51 g. This marginally exceeds the attempted differential of 9·32 g, which confirms that natural selection (as it affects body weight) did not operate, at least not through infertility.

Bearing in mind the sampling errors attached to the generation means, and that the cumulated selection differential was less than 10 g, we must

point out that the residual heritability in the LAU population could have been as high as 10 or perhaps even 20%, without any response being detectable over the period of the experiment. Nevertheless, we do not believe this to be the case; there is little, if any, evidence that additive variance remained in the base population, QLA 29, and that the line at some point prior to that had reached a limit to selection for this reason.

The independent effect of inbreeding (IC)

The mean weights for four generations of sib mating (IC 1–4) are shown also in Table 2. Even though the inbreeding coefficient is 0·59 for the

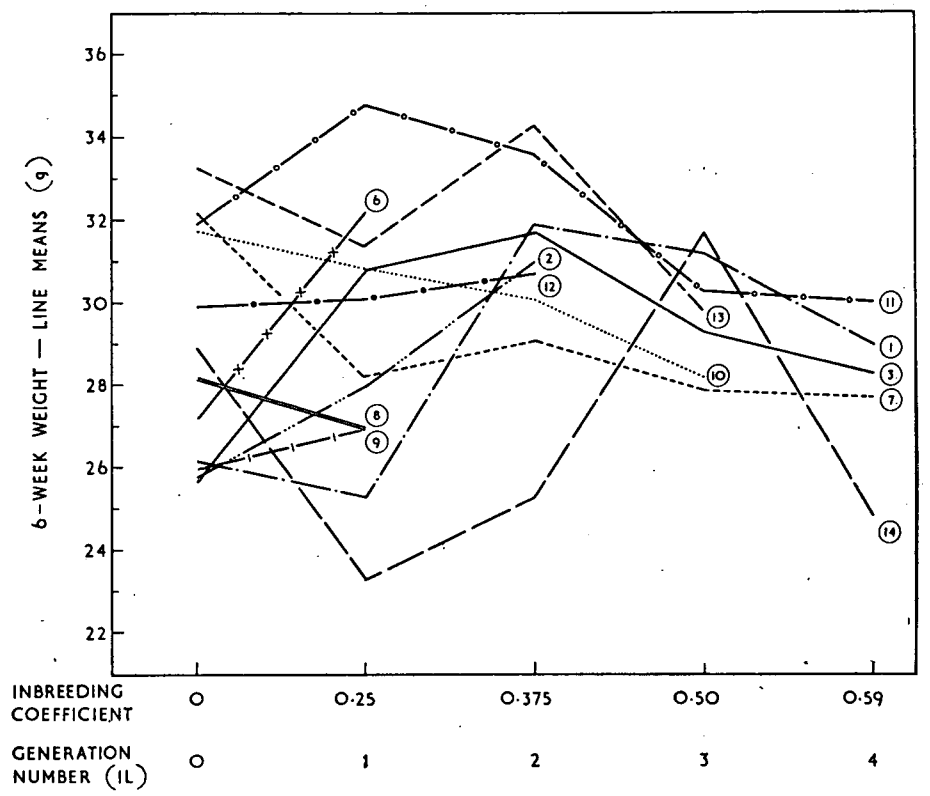


FIG. 3. Mean 6-week weights of the lines inbred without selection (IC), showing when various lines became extinct.

terminal generation, there is no evidence that inbreeding has significantly reduced the mean weight, whether compared to the contemporaneous relaxed population (QLA) or the selected line (LAU).

Only 5 out of the original 16 lines survived to the fourth generation of sib mating. Some lines dropped out at each generation, but as can be seen from Figure 3, the lines that fail to survive are at no time associated with extreme weights, either high or low. However, inbreeding had a marked effect in reducing variation within lines, while the lines themselves became

increasingly differentiated. The results of analyses of variance, on each generation in turn, are shown in Table 3. Because of the small number of lines, and of litters within lines, no meaningful estimates of components of variance could be obtained. It is, then, perhaps fortunate that the qualitative conclusion regarding the differentiation of the lines is so unambiguous. It provides clear evidence that there must have been at least some genetic variance still left in the base population, QLA.

TABLE 3

The results of analyses of variance showing significant differentiation in body weight between inbred lines in the IC stock (inbred without-selection)

Generation no.	1		2		3		4	
	d.f.	MS†	d.f.	MS	d.f.	MS	d.f.	MS
Between lines	11	82·29	8	164·18	6	178·56	4	224·21
Between litters	3	98·35	3	51·03	5	8·36	4	13·79
Within litters	111	12·44	81	6·29	92	7·90	60	8·95

† Mean square between lines significant at 0·01 level in all generations.

The evidence from the preceding two sections may be summarized as follows. Firstly, there is little doubt that QLA, in purely operational terms, had reached its limit to selection at some point prior to the commencement of this study. Secondly, the exhaustion of the additive genetic variance would seem to be a sufficient reason for the limit. Thirdly, the differentiation between lines on inbreeding establishes that fixation was not complete, although it provides no evidence on the number of loci still segregating.

Selection for increased body weight with simultaneous inbreeding and subsequent crossing

Given some residual genetic variance in the plateaued line (QLA), we must now examine whether the scheme of selection with inbreeding (IL) and subsequent crossing, described earlier, can exploit this variance. The main results are summarized in Table 4, which shows also the data from a

TABLE 4

Mean 6-week body weights (g) for various generations of selection under inbreeding (IL) and subsequent crossing (IX), and the partial replicate of the experiment (RIL/RIX)

Generation	Inbreeding		6-week wt		6-week wt of corresponding RIL generations	
	F	F_P†	Mean	SE	Mean	SE
QLA 29	0	0	30·61	0·85	(LAU 32)	
IL 1	0·25	0	28·36	0·55	29·95	0·66
IL 2	0·375	0·25	28·03	0·58	29·70	0·77
IL 3	0·50	0·375	30·17	0·59	27·79	0·71
IX 1	0	0·50	30·11	0·77	28·68	1·46
IX 2	0	0	29·78	0·64	(not measured)	

† F_P is coefficient of inbreeding of the parents of the mice measured.

partial replication (RIL) of the experiment, as outlined previously. The body weights shown in Table 4 should be compared with those in Table 2 for QLA, LAU and IC stocks, IL 1 being contemporaneous with QLA 30 and so forth over successive generations.

There is one abrupt change in weights in Table 4 that should be noted. This increase was unexpected, and its statistical significance was such that it induced us to set up the replicated programme (RIL). In retrospect, however, the increase is less alarming, and it can be explained in several ways. It was not fully appreciated at the time that similar increases were

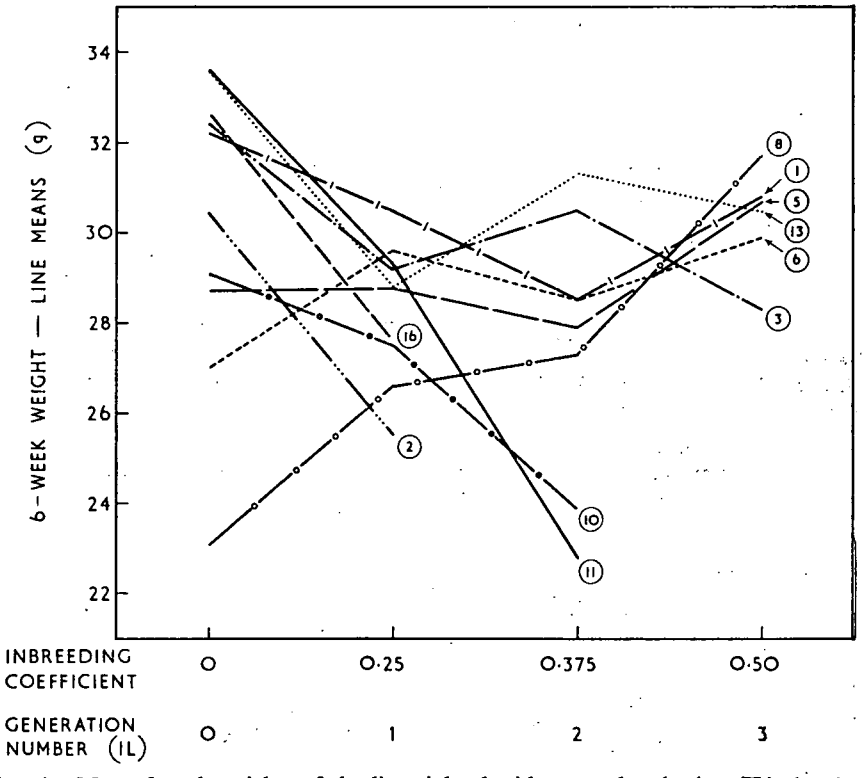


FIG. 4. Mean 6-week weights of the lines inbred with upwards selection (IL) showing when various lines became extinct. Numerical designations of the lines are shown (*cf.* Table 1).

displayed by both the QLA (relaxed) and LAU (selected) lines, and chance environmental factors are probably a sufficient cause. The replicate (RIL) failed to reveal a similar increase at the corresponding but non-contemporaneous stage; in fact, it shows a decline of almost the same magnitude. This would seem to exclude, for example, a systematic effect on body weight through a reduction in litter size, for which the correction may have been inadequate. Such a postulate could not be sustained anyway, as litter size did not change much in the IL programme. The environmental change was therefore much more likely to be some unidentified nutritional or managemental factor, or possibly the amelioration of some subclinical disease. But over and above any environmental shift, the inbred

lines with the lowest weights were continually dropping out of the IL programme (*see* Figure 4), because of the selection. Though the further weights of such lines are obviously indeterminate, it must be supposed that if the selection was at all effective, then this factor could also cause an increase in mean weights as inbreeding progressed.

There is, however, no evidence that the selection was effective. It is true that the depression during the first generation (IL 1) is significant ($-2{\cdot}25$ g$\pm 1{\cdot}01$), and that the eventual recovery of this loss of weight was dismissed above as an environmental shift. But this apart, there is no suggestion that body weights changed at all from the various genetic manoeuvres applied to the stock. No systematic trend is apparent, and

TABLE 5

The results of analyses of variance showing significant differentiation in body weight between inbred lines in the IL (selected upwards with inbreeding and IS (selected downwards with inbreeding) stocks

(A) IL

Generation no.	IL 1		IL 2		IL 3†			
					(*a*)		(*b*)	
	d.f.	MS	d.f.	MS	d.f.	MS	d.f.	MS
Between lines	9	310·05**	8	197·95**	5	209·46**	3	11·43
Between litters	5	61·80	17	4·39	9	58·35	7	70·22
Within litters	118	12·40	119	11·13	103	7·94	79	7·65

† The two analyses for IL 3 refer to: (*a*) all 6 lines that survived to this stage (*b*) the 4 lines used for crossing and which were not longer differentiated.

(B) IS

Generation no.	IS 1		IS 2	
	d.f.	MS	d.f.	MS
Between lines	5	49·84*	3	70·81**
Between litters	0	—	3	13·04
Within litters	32	7·20	44	6·52

* Significant at 0·05 level of probability.
** Significant at 0·01 level of probability.

the terminal point corresponds to the starting weight well within the bounds of sampling error. Further, despite some differences in detail, the replicate agrees with the original within acceptable limits. There is therefore no ambiguity about the main conclusion. The breeding method which Falconer (1971) had successfully employed to improve litter size failed to improve body weight in similar circumstances, and possible reasons for this failure must be examined.

One possible reason for failure has already been discounted. If the fixation of genes had been total, then there would have been no genetic variance of any description left in the stock. Under those circumstances, no method of improvement could possibly work unless it generated new genetic variance. However, the genetic variance in QLA 29 had not been totally exhausted, as shown earlier by the significant differentiation between the IC lines and confirmed by a similar finding in IL (Table 5A). The differ-

entiation among the surviving lines was significant even for the last generation (IL 3). However, only four of the surviving six inbred lines were used for crossing, and for reasons described in an earlier section, the two extreme lines were not used. The differences between the lines actually used for crossing was no longer significant (Table 5A) and this is confirmed by the analysis of the crosses themselves, shown in Table 6.

TABLE 6

Analyses of variance of body weight in the cross (IX I) between the four inbreds (from IL) chosen for crossing

Source	d.f.	MS
Sires	3	20·26
Dams	3	58·24
Sire × dam	4	34·72
Sex	1	232·79
Litters	6	72·64
Residual	111	8·06

Another potential reason why the breeding method might have failed would be the lack of any effective selection differential. We must therefore examine whether the lines contributing to successive generations exceeded in weight those which failed to be represented. This was the case for the first two generations (Table 7) but not for the one just prior to crossing.

TABLE 7

The superiority (g) in body weight of lines contributing to successive generations

Generation	Mean of all lines	Mean of surviving lines	Difference
IL 1	28·36	30·03	$+1·67 \pm 1·26$
IL 2	28·03	28·82	$+0·79 \pm 1·13$
IL 3	30·17	29·01	$-1·16 \pm 0·79$

The accumulated superiority of surviving lines is not very great, and we shall return in the Discussion to the implications of this finding.

Downward selection with inbreeding (IS)

It was explained earlier how this programme was appended to the main experiment. In the event, it yielded a result that clarifies one of the main conclusions. The results are shown in Table 8. Whereas upwards selection failed, downwards selection (from IL 1) yielded a highly significant divergence over a mere two generations. This proves the existence of genetic variation between sib pairs, and confirms that the fixation of genes in the base population was not complete. And as in the other inbreeding studies (IC and IL), there was also significant differentiation between inbred lines in the IS programme (Table 5B).

The selection with inbreeding was thus not so much a failure as ineffective in the upwards direction. The pattern of the divergence between the up and down lines, albeit short term, strongly suggests the segregation of

recessive homozygotes. As soon as they could be identified, downwards selection, in their favour, became effective and the response is clear. It is of course well known that the elimination of such genes by selecting against them is a slow and inefficient process. This explains why the original QLA line, though apparently at its limit, had not reached absolute fixation by the time these studies commenced. While other reasons could easily be adduced for genetic variation in a line at its limit, they are not demanded by data. We need suppose only that fixation of genes affecting body weight would ultimately be achieved, as apparently was the case in another line of mice at its limit for high body weight, described by Roberts (1966b).

TABLE 8

Mean 6-week weights (g) of lines selected upwards (IL) and downwards (IS) during inbreeding, and the divergence between them. Contemporaneous generations of the LAU line (selected without inbreeding) shown for comparison

Generation number	IL		IS		Divergence		*cf.* contemporaneous LAU	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE
1	28·36	0·54					28·16	0·88
2	28·03	0·58	27·76	1·55	0·27	1·65	29·42	0·69
3	30·17	0·59	26·92	0·50	3·25	0·77	30·85	0·68

Generation IS 1 was also IL 1 (*see* text).

DISCUSSION

The following seem to us to be the main experimental findings that have to be reconciled:

(i) The base population, QLA 29, was apparently at its limit to selection.
(ii) There was, nevertheless, significant divergence between inbred lines derived from QLA 29. This occurred in three studies involving inbreeding, (IC, IL and IS).
(iii) Inbreeding alone (IC) failed to yield any significant change in weight, but produced significant differentiation between lines.
(iv) Inbreeding with upwards selection (IL) gave an erratic performance. One change relevant in the present context is the depression in the corrected body weight of 2·25 g$\pm$1·01 on the first generation of inbreeding, before any selection had been applied.
(v) Inbreeding with downwards selection (IS) produced a conspicuous response.

There is therefore no doubt about the main conclusion, namely, that there was some residual genetic variance in the QLA population after it had reached its limit to selection for body weight. The general pattern of the results suggests strongly the segregation of recessives at some loci. We should therefore examine whether there are reasonable sets of values for the variables that affect the population mean, and other parameters, which render the results internally coherent. The relevant variables are the following:

(i) n, the number of segregating recessives,
(ii) q, the mean frequency of such recessives, and
(iii) a, the mean effect of the recessives, defined as the difference between homozygotes.

We should decide beforehand at least the orders of magnitude we are prepared to accept for the different variables. Given that number of segregating genes in an outbred population may not be much more than 20 or so (Roberts 1966a), the number still segregating as the asymptote is approached must be less; we should be reluctant to accept more than 10, and should prefer fewer. Secondly, we consider it unrealistic to postulate gene effects of more than one standard deviation, σ, which in this population was 3 g, and so again, we should prefer smaller effects. However, there may be a basic incompatibility between low numbers and small effects; values that make one variable smaller inevitably render the other larger. Thirdly, we should place an upper limit on the gene frequencies that can be tolerated. In small populations, low values can not apply. While it was being selected, QLA was run on eight pair matings, and on occasions, the number of fertile pairs was no more than five. Thus, any allele present in the stock after such a generation would have a minimum frequency of 0·05, and drift variance would undoubtedly lead to some sharp fluctuations. We must therefore exclude very low frequencies as impossible, and depending on their implications for estimates of various parameters, we should be prepared to accept frequencies of up to 0·20.

It is possible to approach the model from several directions, but it is perhaps best to start with the inbreeding effects, as there is less doubt about the statistical significance of these findings than there is about the other results. The depression on inbreeding is well known to be $2F\Sigma d\,pq$, where F is the inbreeding coefficient, d the deviation of the heterozygote from the mid point between the homozygotes and $q(p = 1-q)$ is the gene frequency. Let us consider the depression in the first generation of IL, which was 2·25 g$\pm$1·01, as noted earlier. The value of F at this point is 0·25; for the postulated recessives, $d = \frac{1}{2}a$, in terms of the earlier definition of a, and Σ can be replaced by n, the number of genes. By equating the expected depression with the observed, we thus have

$$0{\cdot}25\, na(1-q)q = 2{\cdot}25 \pm 1{\cdot}01$$

and the problem is simply to find realistic values of n, a and q that will satisfy the equation to within, say, two standard errors. If we express the gene effect (a) in terms of the phenotypic standard deviation (σ), which was about 3 g, the relationship becomes

$$n\left(\frac{a}{\sigma}\right)(1-q)q = 3{\cdot}0 \pm 1{\cdot}35$$

and any combination of values whose multiple product lies between 0·3 and 5·7 therefore meets the requirement. The upper limits to these values, which we set earlier, gives a product of 1·6, which obviously falls neatly within the required range. Lower values become progressively less probable and it seems we must invoke at least 6 genes, each with an effect of 0·67 σ and at frequencies of 0·2 or so.

Now, the values of these variables suggest substantial genetic effects, and

the next question obviously concerns their compatibility with the lack of response to selection. Unfortunately, our estimates of heritability do not meet the criteria for an adequate test of this question. Falconer (1974), working on the ancestral stock of the mice described here, reported that the realized heritability was 0·37. This value would have been eroded by selection, because of the tendency towards fixation of the genes contributing to body weight. It seems improbable that the residual heritability should exceed 0·10, though the argument at this point becomes totally intuitive.

We have therefore merely examined what the heritability would be, over the range of values earlier found to be compatible with the inbreeding effects, from the formula:

$$h^2 = 2n(a/\sigma)^2q^3(1-q)$$

which enables us to calculate the heritability for different values of the variables. Taking 6 genes with effects of 0·67 σ and at frequencies of 0·2, as we postulated earlier, we find a heritability of 0·034, which in turn is eminently acceptable also. Other values may be explored empirically, and if we exclude constellations giving heritabilities above 0·10, the impression is confirmed that values in the region of those quoted above are the most realistic.

The final question is a straightforward one. Given that number of recessive genes with that order of gene effects, at the frequencies quoted, what would be the improvement in the population mean if such recessives were eliminated? This improvement is equal to naq^2, and by substituting the values just suggested, the answer is found to be 0·48 g. Even if the improvement programme were considered capable of achieving the whole of this potential gain, we still could not have hoped to observe the improvement, given the standard errors attached to the generation means of the IL programme. In other words, the total advance to be obtained—whatever the method of improvement attempted—was insufficient to justify the scale of experimentation that would be necessary to identify the gain at a reasonable level of significance.

No one could rely on the calculations we have presented being arithmetically precise. Nevertheless, they do establish that in lines selected to, or near, their limits, recessive genes of quite large effects, in fair numbers and at almost intermediate frequencies, can still be found segregating. The model lends internal coherence to our rather diverse experimental findings, although the plausibility of the model does not guarantee that it is correct. It has merely been shown that if the model were true, then it would yield a pattern of results similar to that actually observed.

There is no doubt that continued artificial selection would ultimately eliminate the recessives. The selection, however, would become increasingly inefficient and in any real-life situation, with populations of finite size, drift might very well be the final cause of their extinction. Roberts (1966b) reported that in a line with a longer history of continued selection for large size, the fixation of loci affecting body weight seemed to have been complete. The main objective of this study was to explore the generality of Falconer's (1971) method of breaking through a selection limit, by exploiting the residual genetic variance within the line. We found the same general picture of recessives still segregating at the limit, but the consequences of this finding differed. The total gain to be achieved by the elimination of the recessives

from our material was so small that no method of improvement could have worked unless it introduced new genetic variance. Falconer's method presupposes a substantial non-additive component still remaining after selection, and body weight is a character that fails to meet this criterion. We may speculate whether the method might have contributed something useful if it had been applied at an earlier stage in the selection programme. Selection, while still capable of yielding further response, would be becoming increasingly inefficient. At that stage, it might have been profitable to apply the pressure, so to speak, on segregating recessives and expose them by inbreeding. This application of the method might facilitate the terminal stages of selection in traits like body weight where directional dominance is known to be towards large size. However, to justify the switch to inbreeding with continued selection, the final gain would have to be achieved in demonstrably less time. There may be scope for more laboratory experimentation on this point, but its possible application to animal breeding involves also economic factors which are specific to each class of livestock.

ACKNOWLEDGEMENTS

W. K. A-M. is grateful to the British Council and to the Calouste Gulbenkian Foundation for financial support at different times, and to the University of Baghdad for the opportunity to conduct this work. Both authors thank Dr W. G. Hill for his advice on statistical procedures, and Professor D. S. Falconer and Dr Heather G. Pidduck for reading the manuscript.

REFERENCES

FALCONER, D. S. 1971. Improvement of litter size in a strain of mice at a selection limit. *Genet. Res* **17**: 215–235.

FALCONER, D. S. 1974. Replicated selection for body weight in mice. *Genet. Res.* **22**: 291–321.

FALCONER, D. S. and KING, J. W. B. 1953. A study of selection limits in the mouse. *J. Genet.* **51**: 561–581.

HILL, W. G. 1972. Estimation of realized heritability from selection experiments. II. Selection in one direction. *Biometrics* **28**: 767–780.

ROBERTS, R. C. 1966a. The limits to artificial selection for body weight in the mouse. I. The limits attained in earlier experiments. *Genet. Res.* **8**: 347–360.

ROBERTS, R. C. 1966b. The limits to artificial selection for body weight in the mouse. II. The genetic nature of the limits. *Genet. Res.* **8**: 361–375.

ROBERTS, R. C. 1967a. The limits to artificial selection for body weight in the mouse. III. Selection from crosses between previously selected lines. *Genet. Res.* **9**: 73–85.

ROBERTS, R. C. 1967b. The limits to artificial selection for body weight in the mouse. IV. Sources of new genetic variance—irradiation and outcrossing. *Genet. Res.* **9**: 87–98.

(*Received* 15 *February* 1974)

PAPER 11

Maternal effects on body weight in mice selected for large and small size.

Genet. Res. 32, 295-302. 1978.

by

W.K. AL-MURRANI & R.C. ROBERTS.

Genet. Res., Camb. (1978), **32**, *pp.* 295–302
With 2 text-figures
Printed in Great Britain

Maternal effects on body weight in mice selected for large and small size

By W. K. AL-MURRANI* and R. C. ROBERTS†
Institute of Animal Genetics, West Mains Road, Edinburgh EH9 3JN

(*Received* 14 *August* 1978)

SUMMARY

Fertilized eggs were transferred reciprocally between large and small mothers, to study maternal effects on body weight in mice selected for large and small size, respectively. Prenatal maternal effects were not important in our material, but postnatal maternal effects were detectable. The postnatal effects accrued mostly from the inadequacy of small mothers for large offspring; small offspring were largely unaffected by the type of mother. Genetic maternal effects were only of limited importance; maternal influences of environmental origin arose from variation in litter size.

1. INTRODUCTION

Maternal effects on body weight in the mouse have been reviewed by Legates (1972) and by Eisen (1974). Generally, while maternal effects influence juvenile weights, they fade in importance as the mouse approaches maturity. Two main kinds of maternal effects are recognized: prenatal, associated with uterine properties, and postnatal, stemming from lactational or mothering abilities. Some of these effects may be a direct consequence of body size, e.g. large mice may have large uteri and large mammary glands. But there may be genes affecting maternal properties that do not operate through body size. Maternal effects may be either environmental or genetic; some implications of this distinction for biometrical analysis were discussed by Willham (1963) and by Falconer (1964). In addition, there may be interactions between the strain of the mother and the strain of the young, noted particularly by Brumby (1960) and by Mason, Nicholson, Bogart & Krueger (1960).

This paper reports on maternal effects in lines of mice selected for high and low body weight. The question was: To what extent was the difference in body weight attributable to genes directly affecting the character and how far to maternal effects associated with large and small mothers? In the selected lines, the two causes are confounded. To separate them, fertilized eggs were transferred reciprocally between large and small mothers, and the resulting offspring were reared in the litters in which they were born. Thus, large and small mice were

* Present address: College of Veterinary Medicine, University of Baghdad, Al-America, Iraq.
† Agricultural Research Council, Unit of Animal Genetics.

0016-6723/78/2828-7150 $01·00

subjected to the totality of the maternal environment of the other, and their subsequent growth measured.

2. MATERIALS AND METHODS

The mice were taken from generations 33 to 35 of the replicated Q lines, selected for high and low 6-week weight as described by Falconer (1973). Two stocks, one large and one small, were constructed for this study, the stocks being labelled with a colour marker to facilitate the identification of transferred offspring. The large (L) stock was coloured, and comprised samples from four of Falconer's six high lines. The small (S) stock was albino, and comprised samples from three of Falconer's six low lines. The albino gene did not affect body weight, the weighted mean difference (albino minus coloured) in the small parental lines being $-0{\cdot}178 \pm 0{\cdot}225$ g at 6 weeks of age. Each stock was based initially on about 20 matings. The mean 6-week body weights of the L and S stocks, in natural matings over the period of study, were around 30 and 15 g, respectively, in good agreement with the parental Q lines over the same period.

The design is summarized in Fig. 1. Fertilized eggs were taken from L and S females, mated to males of their own stock, and transferred to pseudopregnant females either of their own stock or of the other, reciprocally. The number of eggs transferred was either ten or five. In addition, mixtures of 5L and 5S eggs were

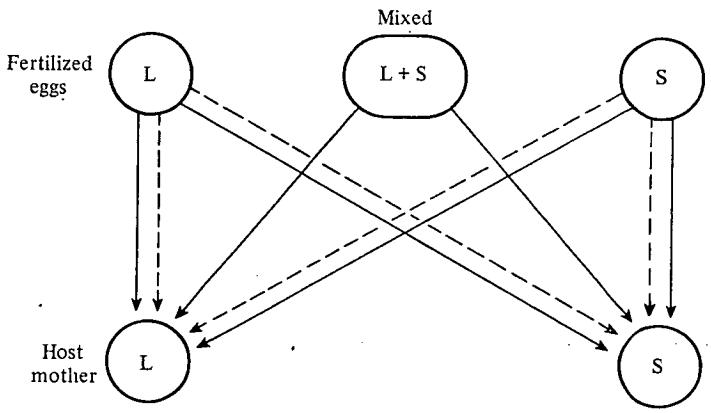


Fig. 1. Diagrammatic representation of the ten groups of transferred eggs. L and S represent large and small genotypes, respectively. Solid lines represent ten eggs transferred, and broken lines five eggs.

transferred into both L and S females. Transfers were conducted as described originally by McLaren & Michie (1956). Natural mating was used throughout, day 0 being when a vaginal plug was found. Most transfers were of $3\frac{1}{2}$-day eggs into $2\frac{1}{2}$-day mothers, with some synchronous transfers at $3\frac{1}{2}$ days to avoid wastage (birth weights were the same in the two groups). Pseudopregnancy was induced by vasectomized males.

The ten groups of transferred eggs (Table 1) are designated by three letters: the

first, T(en) or F(ive), shows the number of eggs transferred, the second, L(arge), S(mall) or M(ixture) shows the genotype of the transferred eggs, and the third letter shows the genotype of the host mother. The number of transfer operations and recovery rates are also shown in Table 1. We hesitate to make any claims for the lower recovery rate from mixed transfers; it may mean nothing more than unavoidable delays during the transfer operations, as eggs had to be collected from different sources.

Table 1. *Number of successful transfer operations and mean number born in each group*

Designation	No. of transfers	Total no. of offspring born	Mean litter size at birth
FLL	12	37	3·08
FSL	9	29	3·22
TLL	11	83	7·55
TSL	7	57	8·14
FLS	28	82	2·93
FSS	15	50	3·33
TLS	8	60	7·50
TSS	12	91	7·58
TML	19	111	5·84
TMS	25	112	4·48
Totals	146	712	

3. RESULTS

As reported also by Brumby (1960), egg transfer *per se* had no effect on subsequent body weight, when comparative data were available. For instance, the TLL group had a mean litter size very close to the natural one of a stock (LAU) used by Al-Murrani & Roberts (1974), whose genetic history was similar; contemporaneous body weights in the two stocks were virtually identical at all ages.

(i) *Body weight differences in a standard maternal environment*

Table 2 shows the body weights of L and S genotypes in blocks where the maternal environment was the same for the two. The objective of a common maternal environment was most clearly achieved in the mixtures, in the same mothers, as this excludes accidental variation in litter size and sampling differences between dams. The difference between L and S within TML may be compared directly with (TLL-TSL), and similarly within TMS with (TLS-TSS). The agreement is generally good, and where any discrepancy is suggested, it falls well short of statistical significance. The difference in growth between L and S was magnified when only five eggs were transferred.

We conclude the following about the growth of L and S genotypes when maternal influences are removed. First, for similar litter sizes, birth weight is virtually unaffected by the genotype of the mother, and is overwhelmingly

a property of the offspring themselves, in these stocks. This excludes genetic prenatal maternal effects, leaving only environmental effects through variation in litter size. The comparisons set out in Table 3 support this with one exception (FLL-FLS); where only five L eggs were transferred, large mothers conferred some advantage. The superiority of L mothers is discussed further in the next section.

The second conclusion is that L genotypes exceed the weight of S genotypes by some 25 % at birth, and this magnifies to about 80 % at 6 weeks. This divergence, however, does not occur uniformly over time. The progeny from the mixed transfers

Table 2. *Body weights (g) at three ages in the different groups (unweighted mean of the two sexes),* ± 1 S.E.

	Mean body weight at:		
Designation	Birth	3 weeks	6 weeks
FLL	2·11 ± 0·05	12·69 ± 0·48	34·97 ± 0·96
FSL	1·47 ± 0·07	9·21 ± 0·25	16·63 ± 0·45
Difference	0·64 ± 0·09	3·48 ± 0·54	18·34 ± 1·06
TLL	1·75 ± 0·03	10·53 ± 0·57	30·62 ± 0·71
TSL	1·45 ± 0·04	8·12 ± 0·57	15·53 ± 0·29
Difference	0·30 ± 0·05	2·41 ± 0·81	15·09 ± 0·77
FLS	1·89 ± 0·04	11·42 ± 0·50	31·45 ± 0·79
FSS	1·55 ± 0·04	8·66 ± 0·41	16·78 ± 0·63
Difference	0·34 ± 0·06	2·76 ± 0·65	14·67 ± 1·01
TLS	1·62 ± 0·07	7·78 ± 1·16	26·97 ± 1·99
TSS	1·42 ± 0·04	6·91 ± 0·32	15·09 ± 0·49
Difference	0·20 ± 0·08	0·87 ± 1·20	11·88 ± 2·05
TML (L)*	1·86 ± 0·05	12·81 ± 0·60	33·03 ± 1·03
(S)	1·50 ± 0·05	9·54 ± 0·43	18·00 ± 0·44
Difference	0·36 ± 0·07	3·27 ± 0·74	15·03 ± 1·12
TMS (L)	1·86 ± 0·04	10·28 ± 0·58	30·54 ± 0·84
(S)	1·46 ± 0·04	8·21 ± 0·38	17·09 ± 0·59
Difference	0·40 ± 0·06	2·07 ± 0·69	13·45 ± 1·03

* L offspring from TML, etc.

were weighed every 3 days, and the L/S ratio of weights is shown in Fig. 2, for L and S mothers separately. The ratio actually declines to 1·15 by 15 days, and only then do the genotypes diverge further. The L genotypes begin to express their superiority before weaning (21 days), possibly due to the earlier eruption of their molar teeth and their ability to eat solid food. Falconer (1973) had reported that most of the divergence between L and S genotypes occurred after weaning. These more detailed data suggest that differential growth starts earlier.

Falconer (1973) reported a L/S ratio of 2·3 at 6 weeks of age. The same ratio for various comparisons from Table 2 ranges from 1·8 to 2·1 when the maternal environment was standardized. Thus, maternal effects may have been responsible for some 10–20 % of the divergence in body weight brought about by the original

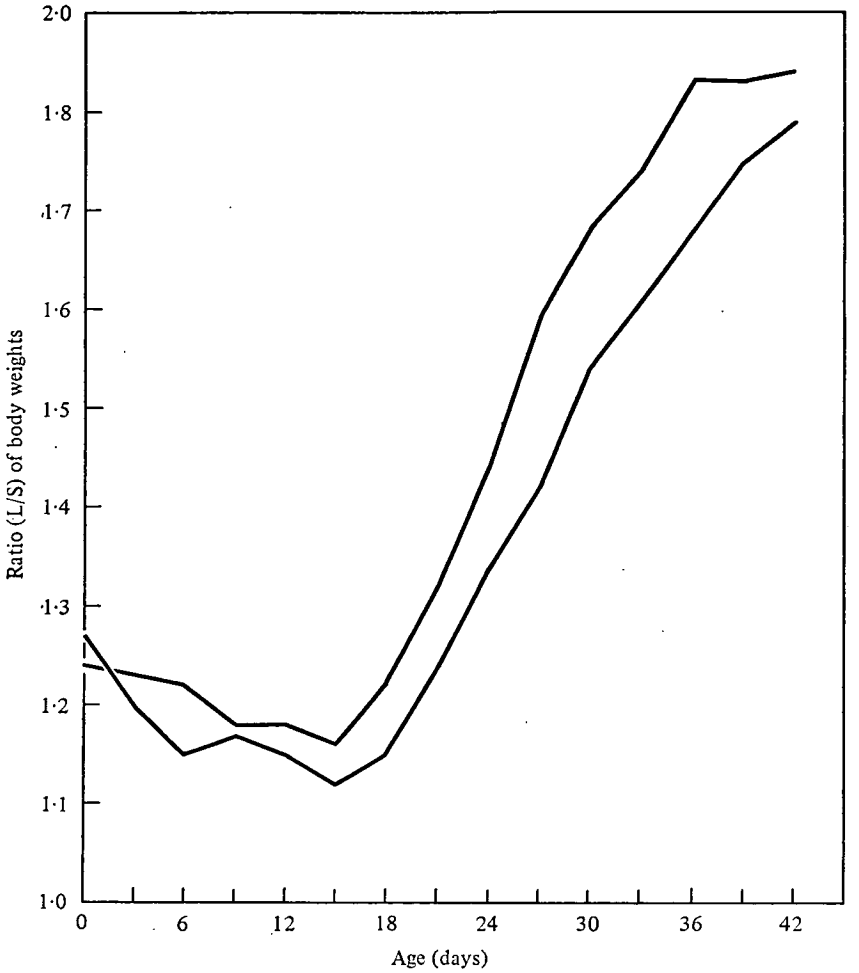


Fig. 2. Ratio (L/S) of body weights of L and S genotypes from mixed litters at different ages; upper line when gestation and suckling by large mothers, lower line by small mothers.

Table 3. *Maternal effects on body weight at various ages. Groups with the same genotype and same number of transferred eggs compared in large and small mothers. Data from Table 2*

	Difference in mean body weight (g) at:		
Comparison	Birth	3 weeks	6 weeks
FLL-FLS	0·22 ± 0·064	1·27 ± 0·69	3·52 ± 1·24
TLL-TLS	0·13 ± 0·076	2·75 ± 1·29	3·65 ± 2·11
FSL-FSS	−0·08 ± 0·081	0·55 ± 0·48	−0·15 ± 0·77
TSL-TSS	0·03 ± 0·057	1·21 ± 0·65	0·44 ± 0·57

selection. These maternal effects were almost wholly postnatal, reflecting lactational performance or some other aspect of maternal care.

(ii) *Maternal influences on growth*

As Table 2 shows, offspring weights were uniformly higher from L mothers, but this was mostly attributable to the depression in weight of L offspring in S mothers. The relevant comparisons are set out in Table 3. The last two rows show that by 6 weeks, S offspring were unaffected by the genotype of the mother, having derived only a small and transient advantage from L mothers at weaning time. L offspring, on the other hand, find S mothers relatively inadequate at all stages (first two rows of Table 3). The postnatal maternal superiority of L mothers, shown in these comparisons, substantiates the conclusion reached earlier. But only L offspring are able to retain this advantage by 6 weeks; it disappears in S offspring.

Table 4. *Relative growth before and after weaning, at 3 weeks of age.* W_0, W_3 *and* W_6 *are body weights at birth, 3 and 6 weeks, respectively. Unweighted means calculated by pooling values from Table 2*

Offspring	Parent	$\frac{W_3-W_0}{W_0}$	$\frac{W_6-W_3}{W_3}$
L	L	5·31	1·75
L	S	4·46	2·06
S	L	5·08	0·87
S	S	4·36	1·07

Table 5. *Regression coefficients of body weight at birth and at 3 weeks on number born alive, and at 6 weeks on number weaned. Data from either five or ten eggs transplanted into each group*

Source	No. of litters	Regression coefficients (g/mouse) of: Birth wt./no. born	3 weeks wt./no. born	6 weeks wt./no. weaned
FLL + TLL	23	−0·069 ± 0·010	−0·27 ± 0·15	−0·53 ± 0·28
FLS + TLS	36	−0·066 ± 0·010	−0·64 ± 0·18	−0·33 ± 0·37
FSL + TSL	27	−0·010 ± 0·010	−0·29 ± 0·08	−0·27 ± 0·12
FSS + TSS	16	−0·023 ± 0·010	−0·39 ± 0·10	−0·41 ± 0·16

The L offspring retain maternal advantages into the postweaning period despite the counterbalancing effects of compensatory growth, whereby growth during any period is inversely related to the proportion of normal growth already achieved, as discussed by Monteiro & Falconer (1966). Table 4 shows the effects of compensatory growth in our material. Preweaning growth is somewhat depressed by S mothers, as noted earlier, but this leads to an increase in relative growth after weaning. The system thus behaves as if it has a built-in correction for maternal effects, though the correction is only a partial one in the case of L offspring.

Maternal effects clearly arise from variation in the litter size, as a result of either

five or ten eggs being transferred. The expected inverse relationship between body weight and litter size (Table 5) is almost uniformly significant for all groups at birth and at weaning time. The absolute difference is, on average, retained until 6 weeks of age, though its significance obviously declines as its relative importance diminishes.

4. DISCUSSION

We conclude that maternal effects are of limited importance in our material. Before birth, they arise only from variation in litter size; during lactation, the same environmental source is detectable. But postnatally, L mothers are superior to S mothers in milk supply, or in some other aspect of maternal care. To the extent that this is a property of the mother's genotype (whether or not it is mediated through body weight), it may be termed a genetic maternal effect. Its contribution to the original selection response was at most 20%, and generally seemed to be somewhat less than this.

Cumulatively, however, maternal effects may be substantial. If we take our extreme comparison from Table 2 (FLL-TLS), the difference in 6-week weight is 8 g, almost 25% of the mean. But if we interpose an intermediate group (FLS or TLL), and split the difference accordingly, we see that litter size alone is responsible for fully half of the difference.

Several of the studies reviewed by Legates (1972) and by Eisen (1974) suggest that maternal effects have their maximum effect around 12–14 days *post partum*, which age coincides with the peak of lactation (Hanrahan & Eisen, 1970). But Monteiro & Falconer (1966) reported that maternal effects increased until 4 weeks of age, 1 week after weaning, and Brumby (1960) reported their persistence even to 12 weeks of age. In our material, some maternal effects were still detectable at 6 weeks of age, but this arose entirely because large offspring had been handicapped by small mothers. Small offspring, on the other hand, showed no residual maternal effects at 6 weeks, and even earlier, had failed to profit to any extent from the superiority of large mothers. These results could be described formally as an interaction between the strain of the mother and the strain of the offspring. Formal descriptions in such terms, however, are not very instructive.

Because of the ubiquity and magnitude of maternal effects, within-litter selection has frequently been favoured when selecting for body weight in the mouse. It avoids some of the complications, even though on other grounds its theoretical advantage is dubious. But even where within-litter selection has been used, as in our material, maternal effects nevertheless accrue. The divergence in body size had not detectably affected uterine performance, but lactational performance had been altered to correspond to the greater postnatal growth of the large lines. Even so, it is arguable whether the improvement in milk supply was adequate for the potential growth of the large lines, if we set as our standard the growth in the reduced litters following the transfer of five eggs only. Further, there was a rapid enhancement of the divergence between large and small mice just before weaning. This could be in part a reflexion of the suboptimal nutrition

of the large offspring up to that time, suppressing the expression of the full genetic difference in body weight and allowing the subsequent influence of compensatory growth.

W.K. A.-M. thanks the British Council and the Calouste Gulbenkian Foundation for financial support, and the University of Baghdad for leave of absence.

REFERENCES

Al-Murrani, W. K. & Roberts, R. C. (1974). Genetic variation in a line of mice selected to its limit for high body weight. *Animal Production* **19**, 273–289.

Brumby, P. J. (1960). The influence of the maternal environment on growth in mice. *Heredity* **14**, 1–18.

Eisen, E. J. (1974). The laboratory mouse as a mammalian model for the genetics of growth. In *Proceedings of 1st World Congress on Genetics Applied to Animal Production*, Madrid, Spain, 1974, vol. 1, 467–492.

Falconer, D. S. (1964). Maternal effects and selection response. In *Genetics Today. Proceedings of XIth International Congress of Genetics*, The Hague, 1963, vol. III, 763–774.

Falconer, D. S. (1973). Replicated selection for body weight in mice. *Genetical Research* **22**, 291–321.

Hanrahan, J. P. & Eisen, E. J. (1970). A lactation curve for mice. *Laboratory Animal Care* **20**, 101–104.

Legates, J. E. (1972). The role of maternal effects in animal breeding. IV. Maternal effects in laboratory species. *Journal of Animal Science* **35**, 1294–1302.

McLaren, A. & Michie, D. (1956). Studies on the transfer of fertilized mouse eggs to uterine foster-mothers. I. Factors affecting the implantation and survival of native and transferred eggs. *Journal of Experimental Biology* **33**, 394–416.

Mason, R. W., Nicholson, H. H., Bogart, R. & Krueger, H. (1960). Predominance of hybrid losses or negative heterosis in mouse crosses. In *Biometrical Genetics* (ed. O. Kempthorne). *International Union of Biological Sciences*, series B (Colloquia), no. 38, 188–109.

Monteiro, L. S. & Falconer, D. S. (1966). Compensatory growth and sexual maturity in mice. *Animal Production* **8**, 179–192.

Willham, R. L. (1963). The covariance between relatives for characters composed of components contributed by related individuals. *Biometrics* **19**, 18–27.

PAPER 12.

Selection for efficiency of feed utilisation in growing mice.

Theoret. & Applied Genet. 59, 129-137.
1981

by

E. YUSKEL, W.G. HILL & R.C. ROBERTS

Theor. Appl. Genet. 59, 129-137 (1981)

Selection for Efficiency of Feed Utilization in Growing Mice

E. Yüksel, W.G. Hill and R.C. Roberts
Institute of Animal Genetics, West Mains Road, Edinburgh (Scotland)

Summary. Selection was practised for improved feed efficiency (gain/feed intake) of mice on two alternative feeding regimes. In one set of lines animals were fed ad libitum, in the other set they were individually fed a fixed amount of feed (about 10% below the control ad libitum intake) which was not changed over generations. For each treatment, a pair of replicate lines (E) were selected on efficiency from 3-5 weeks of age for 8 generations and another pair (L) from 5-7 weeks for 7 generations. A control line was maintained for both E and L lines. In terminal generations mice from each line were tested on each feeding regime, and carcasses of ad libitum fed mice were analysed.

The realized heritability (within families) for efficiency averaged 13%, without much variation over treatments. In the E lines efficiency increased by about 18% of the control mean and in the L lines by about 60%, although absolute changes were small, and responses were similar on the two feeding regimes. Weights at the start of test decreased in the E lines and increased in the L lines; weights at the end of test increased in both.

When tested on the alternative regimes, no interactions were detected for live weights, weight gains or efficiency; selection under fixed intake led to the same increase in appetite as did that under ad libitum.

There were no interactions for carcass composition. Selection for efficiency led to an increase in fatness on both selection regimes and both weight ranges.

Key words: Selection – Mice – Feeding Efficiency – Correlation

Introduction

Efficiency of feed utilization of growing animals depends on the interrelationships among food intake, growth and composition of the gain. These are not simply linear: increased appetite leads to increased growth and a spreading of maintenance costs, but if animals become much fatter, this extra gain may be energetically more demanding and at the same time less desired by the consumer. The associations among the characters also depend on the feeding regime: thus under a scheme whereby all animals are fed the same amount of food, gain and efficiency are completely correlated, but not when appetite is given free expression.

Yüksel (1979) has recently reviewed the genetic inter-relationships among the characters in farm and laboratory animals and we merely summarise his findings. Much of the information comes from selection experiments or breeding programmes for single traits. In all species, growth rate and efficiency are highly correlated, and selection for increased growth rate improves efficiency as a correlated response. While selection for increased growth rate increases both food consumption and efficiency, direct selection for efficiency has uncertain consequences on food consumption, except in the pig, where the two are negatively correlated. In other species, changes in food consumption are usually small and, where they occur, uncertain. Changes in carcass composition also fail to yield a regular pattern. Selection for weight gain may increase fatness, not necessarily at the age of selection but at later ages, because of the increased appetite (Hayes and McCarthy 1976). But where selection has been for increased efficiency, the changes in efficiency were generally greater than where selection was for growth rate alone, both in mice (Sutherland et al. 1970, 1974) and in broilers (Pym and Nicholls 1979; Pym and Solvyns, 1979), with the broiler lines selected for efficiency being leaner. In pigs, commercial selection for increased efficiency and leanness has produced little change in daily live weight gain but increased lean gain and reduced food intake (Smith and Fowler 1978); but since intake can be limited managementally, it has been argued that selection for reduced appetite is pointless (Fowler et al. 1976). In laboratory experiments for increased efficiency of lean gain, there were direct responses but little change in food consumption or gross efficiency (Notter et al. 1976; Gosey 1976). There do not appear to have been experiments comparing the genetic changes in efficiency under the alternatives where appetite was given free expression and where it was not.

0040-5752/81/0059/0129/$ 1.80

In the study reported here, direct and correlated responses to selection for gross efficiency of feed conversion in the growing mouse are examined in relation to two experimental variables: feeding regime and the age range over which efficiency is measured. The feeding regimes were either ad libitum or a fixed amount of food fed to each mouse of each generation, at a level intended to correspond to the mean intake of the base population but actually about 10% lower. These two regimes were used since they might lead to qualitatively different efficiencies: on a fixed amount of food, greater growth, and thus greater efficiency, could arise from a reduction in maintenance requirements, including heat loss, or by achieving a nutritionally less costly body composition. When selected on ad libitum, appetite might also change, not obviously in one direction or the other. The two growth periods were either immediate post-weaning gains, from 3 to 5 weeks of age, or between 5 and 7 weeks, as the mice approached maturity. These two age periods were chosen because sexual maturity occurs around 5 weeks and corresponds with the point of inflexion of the growth curve (Monteiro and Falconer 1966). Growth rates are slower after 5 weeks and if associated with different compositions of the gain, efficiency might reflect different processes.

Materials and Methods

Feeding Regimes

Animals on test were maintained in individual cages. Food intake on the ad libitum regime was measured by using feeding baskets; wastage was treated as if eaten, but little waste was observed. Lines on fixed intake were fed individually every two days on an increasing scale. The amount fed was that consumed by 16 mice (8 of each sex) on ad libitum intake, in a preliminary trial on the unselected base population. In the event, this proved to be about 10% less than the amount consumed by the control lines on ad libitum over the period of the experiment. The amount fed on the fixed intake was adjusted every two days to what was appropriate according to age and sex, except during the first (exploratory) generation of selection, when they were fed the equivalent of one standard deviation of food consumption more. During the selection programme no allowance was made for uneaten food by the animals on fixed intake.

Origin and Designation of Lines

Ten lines of mice were developed and designated as follows. Those selected for efficiency between 3 and 5 weeks of age (early growth) were designated E, while those selected for efficiency between 5 and 7 weeks (late growth) were designated L. A second letter denotes the feeding regime during the selection programme, A for ad libitum and F for a fixed amount of feed. Each of the experimental treatments were replicated. Thus, EA1 and EA2 were the two replicates selected between 3 and 5 weeks on ad libitum feed. Two unselected control lines, EC and LC, were also maintained, with efficiencies measured at ages corresponding to those of the selected lines.

The mice came from roughly equal representations of the six unselected control Q lines (Falconer 1973). From generation 31 of these lines, 2-line and then 4-line crosses were made. From among the 4-line crosses, 28 litters were chosen at random from litters containing at least 4 males and 4 females at weaning (21 days) and with their dams visibly pregnant for a second litter. A male and a female from each litter were assigned at random to each of the 4 E lines (generation 0) and mated to avoid inbreeding. A further 8 pairs (one mouse of either sex from each of 16 different litters) formed the EC (control) line. A similar procedure was applied to the second litters to form the 4 L lines and the LC control, subject to adjustment only when four mice of each sex were not available from some litters.

Selection Programme

Each of the ten lines was subsequently maintained on 8 pair matings, with random mating except for avoidance of close relatives, and within-family selection (at random, for the controls) was practised. Each litter ideally provided 7 young for testing, 3 of one sex and 4 of the other, any numerical deficiencies being made up by extra mice from larger litters. From each litter, one mouse of each sex was selected on its deviation from the family (litter) mean. The weight gain (and the food consumption of those fed ad libitum) was measured for each mouse, the criterion of selection being efficiency (weight gain/food consumed). On the fixed intake, in which uneaten food was also charged to the mouse's account, efficiency ranks identically with weight gain.

Because facilities for individual feeding were limited, the control lines measured only when spare capacity was available at the right time: at generations 4, 8 and 9 in the E lines, and 4, 7 and 8 in the F lines. In retrospect, it would have been desirable to secure a more adequate monitoring of the progress of the selection, rather than concentrate on the final outcome. The presentation of the results will be governed by this limitation.

Selection continued for 8 generations in the E lines, and for 7 in the L lines. In the following generation, but without further selection, samples of mice were taken from the lines selected on each feeding regime and tested on the other, all lines being measured over the appropriate age interval.

Body Composition

After another randomly-mated generation (generation 10 for the E lines and 9 for L), body composition was assessed by chemical analysis on animals, all of which had been fed ad libitum. Each litter from each line supplied 1 female and 1 male chosen at random for dissection and analysis at the starting age (3 weeks for E and 5 weeks for L); this group of 16 mice per line, except for minor losses, were analysed as a bulk sample. At the terminal age, two weeks later, the same litters provided three further mice, two of one sex and one of the other, which were chosen to represent a range of terminal body weights. Three samples per line, one of heavy, one of medium and one of small mice, each comprising 8 animals, were analysed in bulk.

After slaughter, the stomach and intestines were removed, leaving the mesenteric fat, and the carcass was weighed and stored at −20°C. Before analysis the carcass was minced three times, using a mincer plate with 3mm holes. A sample of the mince was freeze-dried for 48 hours to obtain the weight of carcass water. Carcass fat was obtained by ether extraction for 16 hours. Total

nitrogen was determined by the Kjeldahl procedure, and protein estimated as N X 6.25. Ash content was obtained by raising the temperature from 150 to 400°C at the rate of 50°C per hour, followed by holding at 400°C for 16 hours and finally at 600° for a further 7 hours.

Results

1 Responses to Selection

There was a high degree of consistency between males and females in all traits over all generations, so all the results are presented as unweighted means of the two sexes. There was, however, much unexplained variation between generations, as shown for feed efficiency in Figure 1. Therefore all results for the selection lines will be shown as deviations from such contemporaneous control values as were available.

Live weights, weight gains, feed intakes and efficiency are

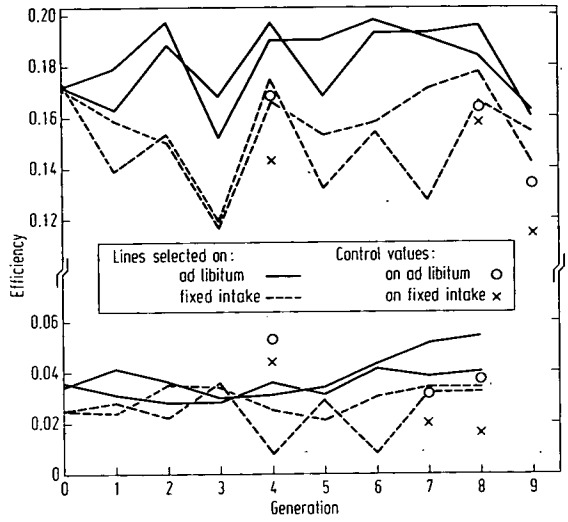


Fig. 1. Mean efficiency of selected lines shown against generation number, E lines above, L lines below, with contemporaneous control values where available.

Table 1. Mean values of live weight, weight gain, feed intake and efficiency of the controls and deviations from controls of lines selected for efficiency between 3 and 5 weeks of age. Efficiency, measured on individuals, is weight gain/feed consumption

Line	Generation	3 wk wt (g)	5 wk wt (g)	Gain (g) 3–5 weeks	Feed Intake (g)	Efficiency
		Selection and measurements on ad libitum feed				
		Control line				
EC (A)	4	10.83	20.59	9.76	57.8	0.168
	8	8.91	17.44	8.53	51.5	0.164
	9	8.06	15.40	7.34	54.3	0.134
		Deviation of selected from control line				
EA1	4	−1.19	−0.77	0.42	−6.0	0.029
	8	−1.40	−0.33	1.09	−2.4	0.032
	9	−0.05	1.11	1.16	−1.1	0.026
EA2	4	−0.73	0.14	0.87	−1.2	0.022
	8	−0.66	−0.17	0.51	−2.4	0.020
	9	0.07	1.51	1.43	−0.6	0.028
		Selection and measurements on fixed feed intake				
		Control line				
EC(F)	4	11.27	18.59	7.32	51.2	0.143
	8	9.57	17.68	8.11	51.2	0.158
	9	8.09	13.98	5.89	51.2	0.115
		Deviation of selected from control line				
EF1	4	−1.11	0.58	1.69	–	0.032
	8	−0.27	0.11	0.38	–	0.008
	9	0.75	2.79	2.05	–	0.039
EF2	4	−1.04	0.15	1.19	–	0.023
	8	−0.50	0.51	1.00	–	0.020
	9	0.23	1.59	1.37	–	0.027
		Average standard errors of deviations from controls (from within line variance)				
	4	0.99	0.74	0.69	2.2	0.014
	8	0.85	0.99	0.54	3.1	0.008
	9	0.66	0.86	0.45	1.6	0.008

Table 2. Mean values of live weight, weight gain, feed intake and efficiency of the controls and deviations from controls and deviations from controls of lines selected for efficiency between 5 and 7 weeks of age. Efficiency, measured on individuals, is weight gain/feed consumption

Line	Generation	5 wk wt (g)	7 wk wt (g)	Gain (g) 5–7 wks	Feed Intake (g)	Efficiency
		Selection and measurements on ad libitum feed				
		Control line				
LC(A)	4	18.32	21.92	3.60	67.6	0.053
	7	19.30	21.27	1.97	63.5	0.031
	8	16.47	19.32	2.85	78.0	0.037
		Deviation of selected from control line				
LA1	4	3.13	1.99	−1.16	10.1	−0.022
	7	0.15	1.96	1.81	9.9	0.020
	8	1.33	2.86	1.53	2.7	0.018
LA2	4	4.19	3.27	−0.92	7.0	−0.017
	7	0.68	1.25	0.56	4.5	0.007
	8	2.78	3.22	0.43	3.6	0.004
		Selection and measurements on fixed feed intake				
LC(F)	4	18.42	21.31	2.89	65.7	0.044
	7	18.91	20.25	1.34	65.7	0.020
	8	16.96	18.01	1.05	65.7	0.016
		Deviation of selected from control line				
LF1	4	2.37	1.13	−1.24	–	−0.019
	7	−0.18	0.75	0.93	–	0.014
	8	1.52	2.68	1.16	–	0.018
LF2	4	2.33	−0.02	−2.35	–	−0.036
	7	−0.56	0.18	0.75	–	0.012
	8	−0.45	0.63	1.08	–	0.016
		Average standard errors of deviations from controls (from within line variance)				
	4	1.03	0.80	0.74	2.1	0.011
	7	1.14	0.79	0.59	2.2	0.008
	8	1.23	0.92	0.61	1.7	0.009

given for the E lines (3-5 weeks) in Table 1 and for the L lines (5-7 weeks) in Table 2. In both, weight gains and efficiencies of the unselected controls were generally higher when the mice were fed ad libitum, while voluntary food intake was also somewhat higher than the fixed amount fed. As a summary, differences between the selected lines and the controls are expressed as a percentage of the control line means in Table 3, with pooling of replicates and the two terminal generations (between which no selection was practised).

In the E lines efficiency improved in all lines in each of the three generations measured, though little progress seems to have been made after the fourth generation (Tables 1 and 3). The improvement in efficiency was very similar on ad libitum and on the fixed food intake, the average improvement being about 18% over the level of the control in the last two generations. Though the selected lines on ad libitum consistently ate less food than the control, the reduction in proportionate terms was very small (except for EA1 in generation 4), so that weight gains over the 3-5 week period of measurement were similar under the two regimes. This greater weight gain tended to be achieved by both a reduction in initial weight and an increase in terminal weight.

Efficiency and weight gains were poorer in the selected L lines than in their control at the fourth generation, when the selected lines also showed much higher initial (5 week) weights (Tables 2, 3). There are not sufficient data to examine a time trend in the LC control line, but it seems more likely that some accident of sampling happened to it rather than simultaneously to the four selected lines. By the final two generations all four L lines had improved efficiencies, LA1 showing the greatest deviation from the control and LA2 the least. Though the improvements in efficiency were small in absolute terms, in proportionate terms the mean increase was 61% over the controls at the end, and much larger than in the E lines. Unlike the E lines, the L lines selected under ad libitum

Table 3. Deviations of selected lines from controls expressed as a percentage of the control mean. Replicates and last two generations pooled

Lines	Generation	Start wt.	End wt.	Gain	Feed intake	Efficiency
		Early lines (3–5 weeks)				
EA	4	−9	−2	7	−6	15
	8 and 9	−6	4	15	−3	18
EF	4	−10	2	20	–	19
	8 and 9	1	9	19	–	19
		Late lines (5–7 weeks)				
LA	4	20	12	−29	13	−37
	7 and 8	7	12	47	8	37
LF	4	13	3	−62	–	−62
	7 and 8	1	6	85	–	86

Table 4. Realized heritabilities (within family) for feed efficiency: total response, calculated as a deviation from controls averaged over the last two generations, divided by cumulative selection differential

Generations	Ad libitum		Fixed	
	Early			
	EA1	EA2	EF1	EF2
8 and 9	0.19 ± 0.04	0.15 ± 0.04	0.13 ± 0.04	0.13 ± 0.04
	Late			
	LA1	LA2	LF1	LF2
7 and 8	0.15 ± 0.03	0.08 ± 0.02	0.11 ± 0.03	0.11 ± 0.03

showed some increase in food intake and their improved efficiency stemmed from even greater weight gains. The greater gains tended to come more from an increase in final weight than reduction in initial weight, particularly under ad libitum.

Realized heritabilities were calculated for the two generations after selection had been stopped by dividing the total response, expressed as a deviation from controls by the cumulative selection differential. Results are shown in Table 4, with standard errors computed by Hill's (1972) method. The estimates are rather consistent and although somewhat higher in the E lines (0.15) than the L lines (0.10), are not significantly so, since only one control line is involved in each case. The overall estimate of realized heritability (h^2_w) is 0.13, which is for selection within full sib families. The intra-class correlation (t) among full sibs for efficiency, averaged for the separate lines over the whole experiment, was about 0.4, and from this the heritability of individual feed efficiency can be calculated as $2(1\text{-}t)h^2_w \sim 0.16$.

Standard errors calculated from the within line variances are given for deviations of line means from the controls in Tables 1 and 2. These are given only as a guide, but cannot be used for significance testing since drift variance is not included. Only for efficiency, on which an estimate of genetic variance could be obtained from the realized responses, could appropriate (although approximate) standard errors be computed as in Table 4.

2 *Tests under the Alternative Feeding Regime*

Generation 9 of the E lines and 8 of the L lines were tested on both feeding regimes, and results are given in Table 5. All lines consumed more food and were somewhat more efficient when fed ad libitum than on the fixed amount fed. But on both feeding regimes, the mice selected on that regime were no better than those selected on the other. The only hint to the contrary comes from line LA1. An analysis of variance, even though using error variances which were too small because drift was ignored, confirmed that line × treatment interactions were non-significant for all traits.

3 *Carcass Analyses*

The results of the analyses of carcasses carried out on animals fed ad libitum at the end of the experiment are summarised in Table 6. The components for the control

Table 5. Comparisons of lines on different diets: control line means and deviations from controls.

	Early lines (3–5 weeks). Generation 9						
	Ad libitum				Fixed intake (51.25 g feed)		
	n	Gain (g)	Feed (g)	Effic.	n	Gain (g)	Effic.
EC	28	7.34	54.3	0.134	29	5.89	0.115
		Deviation from control					
EA1	23	1.16	−1.1	0.026	22	2.03	0.038
EA2	25	1.43	−0.6	0.028	29	1.16	0.026
EF1	19	2.10	3.1	0.030	30	2.05	0.039
EF2	32	1.54	0.9	0.026	31	1.37	0.027
SE[a]		0.46	1.6	0.008		0.44	0.007
	Late lines (5–7 weeks). Generation 8						
	Ad libitum				Fixed intake (65.7 g feed)		
	n	Gain (g)	Feed (g)	Effic.	n	Gain (g)	Effic.
LC	27	2.85	78.0	0.037	28	1.05	0.016
		Deviation from control					
LA1	23	1.53	2.7	0.018	23	1.23	0.019
LA2	23	0.43	3.6	0.004	23	0.54	0.008
LF1	21	0.56	3.5	0.005	23	1.16	0.018
LF2	20	0.31	−0.7	0.005	23	1.08	0.016
SE[a]		0.62	1.7	0.009		0.61	0.009

[a] Average standard error of deviation from controls, calculated from within line variance

lines are shown as percentages of total carcass weight, and values for the selected lines were computed similarly, but are shown as deviations from the controls. Values for the three bulk samples from each line at the terminal weights (5 weeks for E and 7 for L) have been pooled, as there was not a consistent relationship between weight and composition. The standard errors in Table 6 were calculated from an analysis of variance of the sample means, in which the effects of body size and lines of mice within sizes were removed. Strictly, these errors apply only to the samples at the end of the test period, at 5 and 7 weeks respectively for the E and L lines. The errors were applied at the start of the test periods also, as no independent estimates were available.

The two control lines (EC and LC) show good agreement in composition when slaughtered at the same age of 5 weeks. The EC lines increased in fat percentage between 3 and 5 weeks, as expected, but rather unexpectedly, the LC line decreased in fat percentage, as estimated by ether extract, between 5 and 7 weeks.

By far the most striking feature of the carcass results for the selected lines is that, although selected for feed efficiency, they became fatter than the controls at both start and end of test and correspondingly showed a reduced water content. Protein and ash contents also tended to be reduced in the L lines, but not in the E lines at 5 weeks. The replicates within a selection treatment do not appear to resemble each other any more than lines selected on the other treatment. Therefore, just as for traits of the live animal, there appears to be no interaction between feeding regime during selection and carcass composition.

There is, however, some suggestion of a difference between the two ages of the effects of selection on the deposition of fat over the test period. Over the test period, all four E lines became less fat (in percentage terms) than their control but all four L lines put on more fat than their control. Since only one control is involved in each case, the controls may themselves be aberrant, and the declining fat percentage of the LC control from 5 to 7 weeks has already been noted. However, disregarding the controls, the mean fat percentage of the selected E lines increased from 8.4% to 9.7% between 3 and 5 weeks, and that of the selected L lines from 9.2% to 10.0% between 5

Table 6. Carcass composition (%) of control lines, and deviations of selected lines from controls

	Early lines (3–5 weeks). Generation 10							
	3 weeks				5 weeks			
	Water	Fat	Protein	Ash	Water	Fat	Protein	Ash
EC	70.2	6.9	19.0	3.8	69.4	8.7	18.3	3.5
	Deviation from control							
EA1	−1.4	1.8	−0.2	−0.2	−0.8	1.1	−0.4	0.0
EA2	0.6	0.4	−0.6	−0.4	−0.2	−0.8	0.7	0.3
EF1	−1.3	1.3	0.1	−0.1	−1.5	1.1	0.2	0.2
EF2	−1.9	2.7	−0.8	0.0	−3.6	2.4	0.8	0.4
SE[a]	1.5	1.0	0.9	0.5	0.9	0.6	0.5	0.3
	Late lines (5–7 weeks). Generation 9							
	5 weeks				7 weeks			
	Water	Fat	Protein	Ash	Water	Fat	Protein	Ash
LC	69.2	8.4	18.7	3.7	67.8	7.9	20.2	4.1
	Deviation from control							
LA1	−1.0	0.9	−0.2	−0.1	−1.4	1.8	−0.2	−0.1
LA2	−1.1	1.6	0.0	−0.5	−2.0	2.6	−0.4	−0.3
LF1	0.4	0.8	−1.0	−0.2	−1.5	2.4	−0.5	−0.4
LF2	−0.2	0.1	0.4	−0.3	−0.3	1.4	−0.8	−0.4
SE[a]	1.5	1.0	0.9	0.5	0.9	0.6	0.5	0.3

[a] Average SE of deviation from controls estimated from variance among bulk samples within lines

and 7 weeks. Since mean live weight gains were much higher between 3 and 5 than between 5 and 7 weeks (Tables 1, a), it is clear that the E selected lines, though younger, accumulated much more fat than the L lines over their corresponding test periods.

4 *Analyses Within Lines and Generations*

Analyses of variance were conducted within each line and generation to estimate the between and within full-sib family variance and covariance components for weights at start and end of test, gain, feed intake (in A lines) and efficiency. There were no obvious heterogeneities, so results have been pooled over generations and over replicate lines, and are shown in Table 7. This gives the intra-class correlations of each trait, phenotypic correlations between traits and between-family correlations. If maternal effects are ignored, these can be interpreted as one-half the heritability, phenotypic correlations and genetic correlations respectively.

A noticeable feature of the results are the negative phenotypic correlations between efficiency and start weight under ad libitum feeding, the between-family correlations between efficiency and start weight also being negative. The responses to selection (Table 3) were less negative in the E lines and were positive in the L lines. This suggests a major part of these correlations were associated with maternal environment.

Discussion

The realized heritability for efficiency averaged only 13%, and the improvement in efficiency was small in absolute terms. Nevertheless, the E lines exceeded the control means by 20% and the L lines by 60%, after only 7 or 6 generations, respectively, of selection (the last two generations were from random mating). The improvement came almost entirely from increased gain, for food intake on ad libitum changed very little. We may ask therefore whether we could have increased efficiency more by selecting for gain alone. The heritability of gain is usually found to be 2-3 times greater than the value obtained here for efficiency. Estimates of the genetic correlation, obtained from

Table 7. Estimates of intra-class correlation (diagonals), phenotypic correlations (below diagonals) and between-family correlations (above diagonals) from within-line analysis of full sib families over all generations

		E lines				
		3 wk wt	5 wk wt	Gain	Feed intake	Efficiency
3 wk wt	Ad libitum	0.75	0.85	−0.01	0.81	−0.62
	Fixed	0.81	0.59	−0.64	−	+
5 wk wt	Ad libitum	0.67	0.42	0.52	0.84	−0.17
	Fixed	0.47	0.55	0.22	−	+
Gain	Ad libitum	0.06	0.78	0.23	0.27	0.67
	Fixed	−0.54	0.46	0.54	−	+
Feed intake	Ad libitum	0.08	0.79	0.51	0.43	−0.51
	Fixed	−	−	−	−	−
Efficiency	Ad libitum	−0.39	0.33	0.78	−0.11	0.36
	Fixed	+	+	+	−	+
		L lines				
		5 wk wt	7 wk wt	Gain	Feed intake	Efficiency
5 wk wt	Ad libitum	0.50	0.90	−0.66	0.68	−0.75
	Fixed	0.58	0.86	−0.85	−	+
7 wk wt	Ad libitum	0.81	0.38	−0.29	0.75	−0.42
	Fixed	0.75	0.42	−0.45	−	+
Gain	Ad libitum	−0.47	0.12	0.30	−0.18	0.98
	Fixed	−0.72	0.09	0.46	−	+
Feed intake	Ad libitum	0.55	0.65	0.18	0.32	−0.35
	Fixed	−	−	−	−	−
Efficiency	Ad libitum	−0.57	0.00	0.98	0.03	0.26
	Fixed	+	+	+	−	+

No result on fixed intake; + efficiency ≡ gain on fixed intake;
typical SE: intra-class correlation 0.05; phenotypic correlation 0.03, between-family correlation 0.10

full sib families in this study, were 0.67 for the E lines and nearly 1.0 for the L lines (Table 7). Even the lower values suggest that efficiency would have been improved at least as much by selecting for gain alone, while the higher values would predict selection for gain to be much more effective. Broiler breeders have long taken this simple line, avoiding the extra labour of weighing feed as well. Some early pig experiments (Dickerson and Grimes 1947) came to the same conclusion. More recently, however, a broiler experiment (Pym and Nichols 1979) and some pig work (Smith et al. 1962; Vogt et al. 1963; Park 1965) all suggested that direct selection on efficiency is preferable, if that is the trait to be improved. Our conclusion that efficiency would have changed more by selecting for gain should be qualified; we used a short feeding period of two weeks, and some mice may have had difficulty adapting to single cages.

Two features of the data were unexpected. First, mice selected for improved efficiency became fatter, as was also found in similar circumstances, selecting mice for gain on a fixed intake, by McPhee et al. (1980). This does not accord with the simple view that the energetic cost of the accretion of lean tissue is less than that of fat, as a result of the inclusion of so much water in lean. Nor does it accord with pig experience, noted earlier. However, the customary difficulties of translating across species apart, the bioenergetic arguments are complicated. Webster (1977) points out that perhaps some 70% of a growing animal's energetic input is dissipated as heat. While one source of such heat will be the chemical reactions involved in protein synthesis, it seems likely that the differential demands of laying down lean and fat may account for only a part, perhaps a small part, of the total energetic input. The alternative outlets for energy may have swamped the system. As one example, since the mice selected for efficiency tended to be smaller at the start of the test period, their maintenance requirements associated with protein turnover may have been less. If this were the case, selection over a fixed age period (as was done here) may not be directly comparable to selection over a fixed

weight range, as is frequently done with domestic livestock. The selected mice also tended to be fatter at the start of the test period, possibly leading to a reduced maintenance requirement while on test. This observation poses a cautionary note: the effects of selection for efficiency cannot be fully assessed without monitoring changes in metabolic demands outside the test period.

The second unexpected feature of the experiment was the total lack of interaction between feeding regime and the response to selection. This differs from the results of Hetzel (1978), who selected mice for gain both under ad libitum and on a fixed intake. The fixed intake part of his experiment is identical to that part of ours, but the two characters (gain and efficiency) on ad libium are not directly comparable. Hetzel found an interaction between his selection responses and feeding regime: weight gain on each feeding regime was most improved by selection on that regime. The food intake of his line selected on a fixed amount was marginally decreased when tested on ad libitum and neither was the fat percentage of that line significantly altered on either feeding regime. In our case, neither appetite nor carcass composition differed between the two selection methods. The question therefore shifts: why did appetite not change when it was given a free role when selecting on ad libitum feeding? Perhaps the first point to note is that except for the pig, the connection between appetite and efficiency is not very clear (Yüksel 1979). Even so, if there is any genetic variation in efficiency, animals that secured the same weight gain on less food would be selected, and it was precisely this concern that prompted us to introduce the two feeding methods when designing the experiment. However, we probably did not entirely exclude variation in appetite under the fixed regime, since some mice may have failed to eat all of the feed offered but were charged with it anyway. Subjectively, we were not aware of extensive refusals, but they were not measured. But if they occurred, this would be another source of interaction rather than an explanation of a lack of interaction. To the extent that our results may be generalized, the debate among pig breeders about optimal feeding schemes under test might prove to be superfluous.

Acknowledgement

We are much indebted to the Edinburgh School of Agriculture for conducting the carcass analysis.

Literature

Dickerson, G.E.; Grimes, J.C. (1947): Effectiveness of selection for efficiency of gain in Duroc swine. J. Anim. Sci. 6, 265-287

Falconer, D.S. (1973): Replicated selection for body weight in mice. Genet. Res. 22, 291-321

Fowler, V.R.; Bichard, M.; Pease, A.H.R. (1976): Objectives in pig breeding. Anim. Prod. 23, 365-387

Gosey, J.A. (1976): Selection for efficiency of lean tissue growth in rats. Diss. Abstr. B37, pp. 1987

Hayes, J.F.; McCarthy, J.C. (1976): The effects of selection at different ages for high and low body weight on the pattern of fat deposition in mice. Genet. Res. 27, 389-433

Hetzel, D.J.S. (1978): Genetic studies of growth and body composition in mice. Ph.D. Thesis, University of Sydney

Hill, W.G. (1972): Estimation of realized heritabilities from selection experiments. 2: Selection in one direction. Biometrics 28, 767-780

McPhee, C.P.; Trappett, P.C.; Neil, A.R.; Duncalfe, F. (1980): Changes in growth, appetite, food conversion efficiency and body composition in mice selected for post-weaning weight gain on restriced feeding. Theor. Appl. Genet. 57, 49-56

Monteiro, L.S.; Falconer, D.S. (1966): Compensatory growth and sexual maturity in mice. Anim. Prod. 8, 179-192

Notter, D.R.; Dickerson, G.E.; Deshazer, J.A. (1976): Selection for rate and efficiency of lean gain in the rat. Genetics 84, 125-144

Park, Y.I. (1965): Age-constant feed efficiency of pigs. J. Anim. Sci. 24, 819-822

Pym, R.A.E.; Nicholls, P.J. (1979): Selection for food conversion in broilers. 1. Direct and correlated responses to selection for body weight gain, food consumption and food conversion ratio. Brit. Poult. Sci. 20, 73-86

Pym, R.A.E.; Solvyns, A.J. (1979): Selection for food conversion in broilers. 2: Body composition of birds selected for increased body weight gain, food consumption and food conversion. Brit. Poult. Sci. 20, 87-97

Smith, C.; Fowler, V.R. (1978): The importance of selection criteria and feeding regimes in the selection and improvement of pigs. Livest. Prod. Sci. 5, 415-423

Smith, C.; King, J.W.B.; Gilbert, N. (1962): Genetic parameters of British Large White bacon pigs. Anim. Prod. 4, 128-143

Sutherland, T.M.; Biondini, P.E.; Haverland, L.H.; Pettus, D.; Owen, W.B. (1970): Selection for rate of gain, appetite and efficiency of feed utilization in mice. J. Anim. Sci. 31, 1049-1057

Sutherland, T.M.; Biondini, P.E.; Ward, G.M. (1974): Selection for growth rate, feed efficiency and body composition in mice. Genetics 78, 525-540

Vogt, D.W.; Comstock, R.E.; Rempel, W.E. (1963): Genetic correlations between some economically important traits in swine. J. Anim. Sci. 22, 214-217

Webster, A.J.F. (1977): Selection for leanness and the energetic efficiency of growth in meat animals. Proc. Nutr. Soc. 36, 53-59

Yüksel, E. (1979): Genetic aspects of the efficiency of food utilisation in some farm and laboratory animals. Anim. Breed. Abstr. 47, 499-504

Received May 2, 1980
Accepted September 29, 1980
Communicated by A. Robertson

Dr. E. Yüksel
F.Ü. Fen Fakultesi
Zooloji Bolumu
Elazig (Turkey)

Dr. W.G. Hill
Dr. R.C. Roberts
Institute of Animal Genetics*
West Mains Road
Edinburgh EH9 3JN (Scotland)

Book Reviews

Hare, W.C.D.; Singh, E.L.: Cytogenetics in Animal Reproduction. England: Commonwealth Agricultural Bureaux 1979. 96 pp., 42 figs., Hard bound £ 13,20.

'Cytogenetics in Animal Reproduction' by W.C.D. Hare and Elizabeth Singh is the first comprehensive review of cytogenetics and its role in reproduction of domestic animals that has been published.

The first chapters are devoted to defining cytogenetics and explaining fundamental concepts. Different causes of reproductive failure and their cytogenetical background are systematically described and comprise the most central part of the book. For each cause of reproductive failure there is a key for diagnosis. There are also chapters on prenatal cytogenetic studies and animal hybrids. The book is rounded-off with descriptions of culture and preparation techniques as well as the most important banding and photographic techniques used in the authors' laboratory. Forty-two different illustrations – schematic drawings and karyotypes of different domestic animals - are compiled in the final section of the book, together with a very complete reference list.

This book is a very good introduction for the beginner in that it clarifies, in a logical and pedagogic manner, many fundamental concepts in addition to giving advice in different matters. At the same time, it is of great interest to the more advanced reader since it reviews the field carefully and contains a very extensive and complete reference list. Although there are a few printing errors, and as well the quality of some microphotos for different reasons is not very high, the book can be highly recommended. The book can be expected to fill a great need in cytogenetics of domestic animals, not only for students of veterinary medicine and animal breeding, but also for research workers and laymen.

I. Gustavsson, Uppsala

Vorontsov, N.N., Van Brink, J.M. (eds.): Animal Genetics and Evolution. Selected papers of the XIV International Congress of Genetics, August 21-30, 1978, Moscow. The Hague: Dr. W. Junk, B.V. Publishers 1980. 382 pp., 243 figs., 72 tabs. Hard bound Dfl 195,–.

This book is largely a collection of case studies of evolutionary events and trends as indicated by cytological studies. Within this generally creditable volume of work is a lot of speculation as to modes of evolution which should be of interest to those involved in this field. There is also some valuable reading for the quantitative geneticist who feels that his classical guidelines are not of universal validity. One gains the impression that duplication of genome segments and other such phenomena might significantly augment allelic segregation in giving rise to genetic variation between individuals. An example of response to selection in a highly inbred population indicates a significant role played by novel mutations. For the causal reader there is too much overlap in subject material, but, for those with a deeper interest this book could prove to be a standard collection.

H. Skjervold, Ås–N.L.H.

F.J. Ayala, J.A. Kiger, jr.: Modern Genetics. Menlo Park, Calif.: Benjamin/Cumming Publ. Co. 1980. 844 pp., 447 figs., many tabs. Hard bound $ 21.95.

'Modern Genetics' is a textbook intended to be used as an introduction into this fascinating science. The three fundamental features of genes: transmission, expression, and change, are respectively represented in the three main parts of 'Modern Genetics'. Part I deals with the organisation and replication of genetic material, including the Mendelian laws, the nature of genetic material, the eukaryotic and the bacterial genome, and also DNA replication, repair and recombination. Part II treats the expression of the genetic materials. These topics cover a range from the genetic code and genetic function to information transfer in cells and regulation of gene expression in both prokaryotes and eukaryotes. Part III is devoted to the evolution of genetic material. Population genetics and evolutionary genetics, which usually have secondary importance in compendiums of genetics, are treated in greater depth here, partially due to the fact that F. Ayala is an outstanding expert on this field.

At the end of each chapter the student will find problems which should facilitate his understanding and also provide new information.

A glossary and an extensive bibliography are also included. The concepts and methods of statistics included in the appendix are great aids in understanding the text. That this compendium is really up-to-date is shown by the fact that even literature from 1980 is cited. It is noteworthy that the book is excellently illustrated-this enhances the clarity of the text. The book is wholeheartedly recommended to all students of genetics.

Tröbner, Halle

PAPER 13.

Selection for total weaning weight in the mouse, and its implications for domestic livestock.

Z. Tierzucht. Zuchtungsbiol. 99, 222-231.
1982.

by

D.E. STEANE & R.C. ROBERTS

Sonderdruck aus Zeitschrift für Tierzüchtung und Züchtungsbiologie
Bd. 99 (1982), H. 3, S. 222-231

VERLAG PAUL PAREY · HAMBURG 1 · SPITALERSTRASSE 12

ARC Unit of Animal Genetics, Institute of Animal Genetics, Edinburgh, Scotland

Selection for total weaning weight in the mouse, and its implications for domestic livestock

By D. E. STEANE and R. C. ROBERTS

Ms. received 15. 10. 1981

Introduction

Productivity of the female has long been of interest and of importance in animal production. The importance of female productivity has become greater as more intensification has occurred with the increase in the cost of maintaining females.

In general, the measure of productivity has been variously regarded as the total output either per parity, per annum or per lifetime, each definition successively complicating the measurement. Total litter weight at weaning has been used as a measure of female productivity both for pigs and sheep, and the recording of such measurements has been part of many national recording schemes for many years.

The mouse has been used as an experimental model for farm species, as has been discussed on several occasions (eg ROBERTS 1965; EISEN 1974). Much experimental work on maternal performance and, in particular, its prenatal and postnatal genetic control has been carried out but there has been little work done on the total weight of the litter at weaning despite this being the measurement of major importance in livestock production.

A selection experiment for litter weight (by N. BATEMAN) was reported by FALCONER (1955). No progress was made in increasing 12 day litter weight using a standardised litter size. In 1963, DALTON and BYWATER reported an experiment in which lines were selected for either litter size or litter weight at weaning, on each of two diets, with lines being selected both upwards and downwards. The authors reported that there was no response over 14 generations. FRAHM and BROWN (1975) selected for increased preweaning and postweaning gain in mice and likewise found no change in numbers weaned over the generations, but EISEN et al (1970) succeeded in changing 12-day litter weight by selection and estimated the heritiability to be 0.25. This estimate was higher than those reported by LEGATES and FARTHING (1962), who gave realised heritability estimates of 0.04–0.18, and by ROBISON et al (1972) who found a realised heritability of 0.08, in both cases also for 12-day litter weight.

There have been many studies concerning individual components of total weaning weight and the relationships between various components. EISEN (1973) concluded that there was a need either to standardise or eliminate post natal maternal environmental effects. However the relevance of much of this work in terms of farm animals is questionable, since the consequences of large increases in adult body size (which were common in laboratory experiments) have far reaching effects.

Total productivity at weaning is a practical measure of major concern in farm species. The weight of the dam's litter at weaning includes gain due to the consumption of food other than dam's milk – indeed, in pigs there appears to be an inverse relationship between milk availability and feed consumption (BARBER et al, 1955). Whilst EISEN (1973) suggested litter

U.S. Copyright Clearance Center Code Statement: 0044–3581/82/9903–0222 $ 2.50/0
Z. Tierzüchtg. Züchtgsbiol. 99 (1982) 222–231

ISSN 0044–3581/InterCode: ZTZBAS

standardisation, this is not totally feasible in pigs although standardising the environment can be achieved to a degree, as shown by OWEN et al (1978) with sheep by using a ‚complete diet'. Total weaning weight as a composite may be affected in different physiological ways. For instance, the separate effects of ovulation rate and embryonic mortality or the number of young were discussed by FALCONER (1960a, 1963). Genetic work on the trait of total litter weight at weaning has given inconclusive answers. There is a clear need to study this problem further in view of its importance in farm species.

Materials and methods

Experimental procedures

The stock used was a cross of four lines previously selected to their limits for body weight at six weeks of age (see ROBERTS 1967 for details). Fifteen unrelated litters were randomly selected from the 4-line cross, the only requirement being that the litter contained a minimum of two males and four females. Within each litter, four females and two males were selected on the basis of their six week weight. Each full sibship of four females was then mated to an unrelated pair of full sib males in two harems (see Figure 1 for mating design).

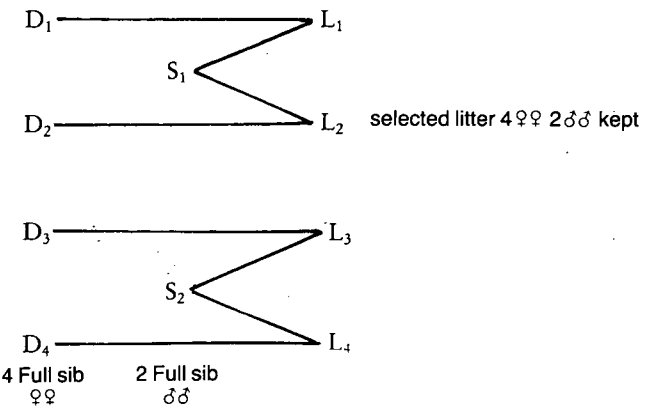


Fig. 1. Mating design

This scheme was repeated over subsequent generations. Matings of first cousins or closer were avoided, as were reciprocal matings. Where possible, the allocation of females to each male was done in a systematic manner to balance out the relative weights of females within the selected group. The availability of four females and two males per selected litter remained as the ideal but obviously one that was not uniformly achieved in practice. Where numbers fell short, the next best litter was taken.

In each generation, about 20% of litters were discarded because they were late born and would unduly prolong the generation interval. Selection was based on Total Weaning Weight (TWW). The litter with the highest weight within a family of full sib females was selected, unless by six weeks (age of mating) there were insufficient mice alive. In generation 13, insufficient litters were reared to allow any selection to take place so that this generation has been excluded from the analysis. Unfortunately, a contemporaneous control population was not available until generation five of the experiment. In the analysis, the control generation contemporaneous with the 13th of the selection line has also been excluded. The selection programme was discontinued after generation 20.

Analysis of the experiment

Two separate analysis were conducted – one of the effect of selection over all generations and the other of the analysis of variance within generations. Generation means were calculated as the mean of family means. The selection differential was calculated on a within family basis.

Analysis of generation means

The generation means were regressed on cumulative selection differential. The results were analysed both over the whole experiment (generations 0–20) and for the period covered by a control line (generations 5–20). For this latter period, deviations from the control were analysed as well as generation means. The component traits Numbers weaned (N) and Average individual weaning weight (W) were also studied in a similar manner.

Analysis within generations

Within generation analysis of variance was carried out on the basis of the model:

$$X_{ijkl} = u + g_i + f_{ij} + m_{ijk} + e_{ijkl}$$

where X_{ijkl} = TWW of the litter of the lth dam
u = overall mean
g_i = effect of the lth generation ($i = 0_i, \ldots . 20$)
f_{ij} = effect of the jth family of full sib dams and the sires of the litters within the ith generation ($j = 1, \ldots . . .15$)
m_{ijk} = effect of the kth male used to sire the litter within the jth family ($k = 1, 2$)
e_{ijkl} = effect of the lth dam within the kth sire ($1 = 1, 2$)

However, since analysis of variance failed to show any effect of the male on any of the traits examined, the model simplifies to

$$X_{ijk} = u + g_i + f_{ij} + e_{ijk}$$

and all the results will be presented in these terms.

There was a positive linear association between the means and variances for the generation means which was successfully removed by a transformation to square roots; subsequent analyses were on the transformed data. The analysis was carried out using Harvey's Least Squares Mean Program (Harvey, 1977) on the square root of TWW, Number Born (B), Number weaned (N) and Average individual weaning weight (W).

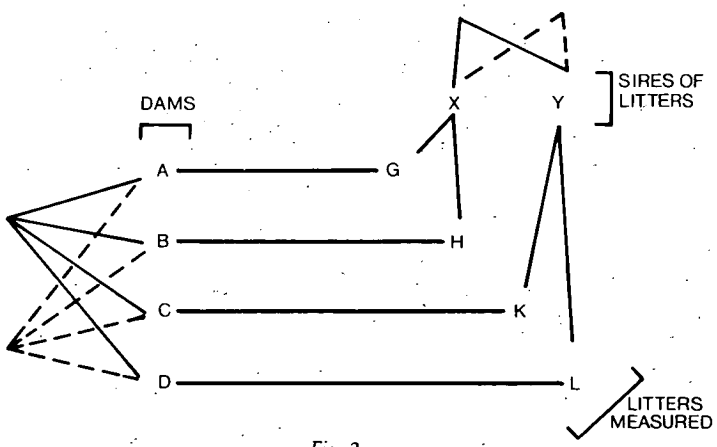


Fig. 2.

Genetic Model

The genetic model considered the trait under selection (TWW) as a trait of the dam but nevertheless the measurement of TWW was on the dam's progeny and, inevitably received a genetic contribution from their sire. The model was further complicated by the relationships between the various levels of the hierarchical structure. The coefficient relationships (R) can be calculated from Figure 2. The model ignores any inbreeding which has occurred previous to the matings considered.

$R_{AB} = R_{AC} = R_{XY} = \frac{1}{2}$ (Full sibs)

$R_{GH} = R_{KL} = \left(\frac{1}{2}\right)^3 + \frac{1}{4} = \frac{3}{8}$ (Paternal half sibs, maternal first cousins)

$R_{GK} = R_{GL} = R_{HK} = R_{HL} = 4 \times \left(\frac{1}{2}\right)^4 = \frac{1}{4}$ (Double first cousins)

The contributions to the Total litter weight can be considered as follows. There are two contributions possible through the dam, one being its direct contribution of genes for growth of the progeny (A_D), the other being through the dam's own genes for maternal (or nursing) ability (A_M). The sire of the litter can contribute additive genetic variance for growth in the progeny. There are two environmental effects, one a general one (E) and the other, the maternal environment (M_E). The components can be summarised using the relationships (R) as calculated (R_P for population = 1).

Variance of individuals within sire of litter $\sigma_I^2 = \sigma_E^2 + (R_P - R_{GH})\,\sigma_{A_D}^2 + (R_P - R_{AB})\,\sigma_{A_M}^2$

Variance of sire of litter within family $\sigma_S^2 = (R_{GH} - R_{GK})\,\sigma_{A_D}^2 + (R_{AB} - R_{AC})\,\sigma_{A_M}^2$

Variance of family means $\sigma_F^2 = (R_{GK})\,\sigma_{A_D}^2 + (R_{AC})\,\sigma_{A_M}^2 + \sigma_{M_E}^2$

The proportions contributed are summarised in Table 1.

Table 1.

	Coefficients of		
	A_D	A_M	M_E
Between family (σ_F^2)	$\frac{1}{4}$	$\frac{1}{2}$	1
Between sire of litter (σ_S^2	$\frac{1}{8}$	0	
Between individual (σ_I^2)	$\frac{5}{8}$	$\frac{1}{2}$	

The maternal environmental effect was ignored in this analysis.

The analysis considered TWW as the major trait but this can be split into several components. The simplest split is with number weaned (N) and average weight of a mouse (W). The initial analysis considered these traits (and numbers born) only on litters for which a weaning weight had been recorded – 735 litters in all. A further analysis included data from the remaining litters (1047 records).

Results

The means for each generation for TWW, N and W for the selected and the control lines are shown in Figures 3 (TWW) and 4 (N and W).
Results over generations

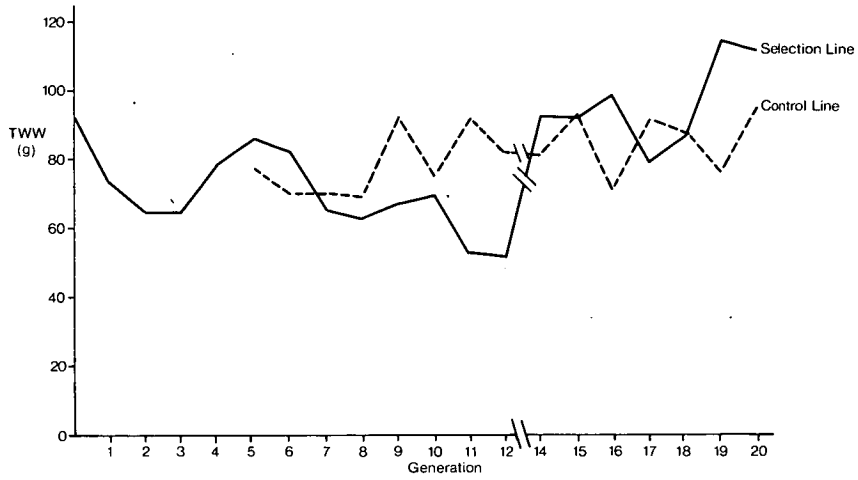


Fig. 3. TWW for Selected and Control Lines

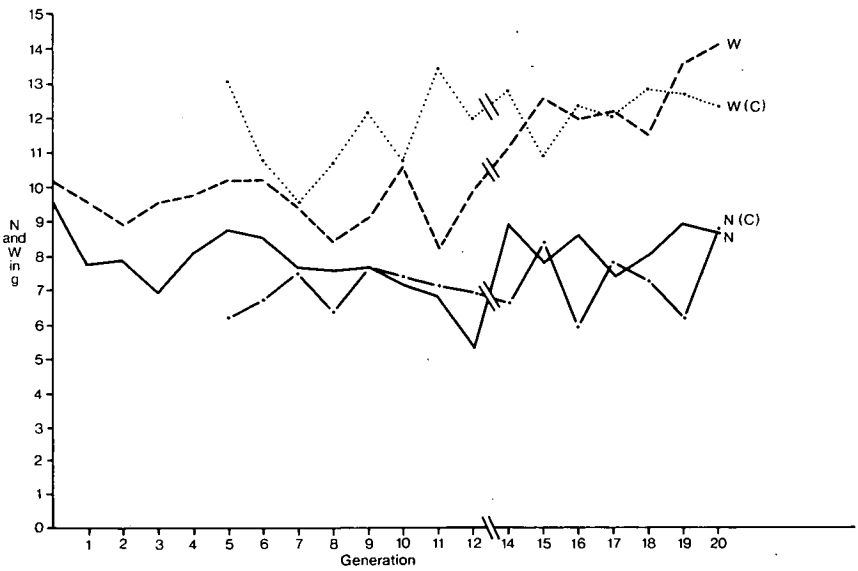


Fig. 4. Number Weaned (N) and Average Individual Weaning Weight (W) by generation

The cumulative selection differential in standard units was 6.554 for TWW on the transformed data, 4.362 for N and 1.948 for W. Standardisation of the untransformed data is not valid, given the linearity between mean and variance. The correlations between the selection differentials were 0.997 (TWW and N), 0.963 (TWW and W) and 0.942 (N and W). The regression of generation means on cumulative selection differential (TWW) are shown in Table 2.

Table 2. **Regression analysis**

Regression	Trait	Generations 0–20	Generations 5–20
Generation mean on Cum. Sel. Diff.	TWW	0.116	0.269
Subtrait means on Cum. Sel. Diff. for TWW	N	−0.008	0.008
	W	0.082	0.156
Deviation from control on Cum. Sel. Diff.	TWW		0.141

None of these regressions are significantly different from zero

As can be seen from figures 3 and 4, there appears to be little or no response with the exception of the last two generations. The non-significant regression in Table 2 confirms this. However, the regression for total weaning weight is still the best method of estimating the heritability. Since selection of both sexes was based solely on the dams own record, the regression has to be doubled. The estimates so obtained range from 23% to 54%, depending on which data are considered, while that based on deviation from control is 28%. Nevertheless, the errors attached to these estimates are such that they are all compatible with zero, and they should be viewed accordingly.

Genetic analysis within generations

Table 3 summarises the expected mean squares, and the coefficients of the components indicate the extent to which inbalance in practice failed to meet the design, which would have led to $2\sigma_S^2$ and $4\sigma_I^2$.

Table 3. **Expected Mean Squares**

Between families within generation	σ_I^2	+	$1.66\,\sigma_S^2$	+	$2.91\,\sigma_F^2$
Between sires of litters within family	σ_I^2	+	$1.60\,\sigma_S^2$		
Between individuals within mates	σ_I^2				

where I = individual, S = sire of litter, F = family

Table 4. **Estimates of heritability (on diagonal); phenotypic correlations (above diagonal) and genetic correlations (below diagonal). Based on all litters weaned (n = 735)**

TWW	0.28 ± 0.09	0.32 ± 0.11	0.65 ± 0.07	0.34 ± 0.11
No Born	0.78 ± 0.20	0.28 ± 0.09	0.79 ± 0.05	−0.63 ± 0.07
No Weaned	0.69 ± 0.17	0.97 ± 0.09	0.18 ± 0.09	−0.47 ± 0.10
Average Weaning	0.45 ± 0.21	−0.21 ± 0.32	−0.33 ± 0.39	0.28 ± 0.09

The estimates of heritability, phenotypic and genetic correlations obtained from the analysis of variance are shown in Table 4. The heritability of the main trait (TWW) is 25%, and for the subtraits 18% to 28% for N and W, respectively, indicating that useful progress is possible. The heritability estimate for total weaning weight agrees well with the regression estimates of realised heritability. Both N and W are favourably correlated with TWW but there is a negative correlation both phenotypically (−0.47) and genetically (−0.33) between N and W.

Discussion

The regression analysis over generations provides an estimate of the realised heritability of TWW. The estimates based on generation means vary from 0.23 to 0.58. These estimates are subject to some doubt since the regressions are not significantly different from zero and if

reference is made to figures 3 and 4 it can be seen that essentially all of the apparent progress occurs in the final two generations. Indeed, if the regressions are calculated from generations 0 to 18 the regression is 0.028, compared to 0.116 when the last two generations are included.

The estimates, as given, are based on similar analysis to those reported by DALTON and BYWATER (1963) and by FRAHM Brown (1975). Both these reports concern selection for Total weaning weight and whilst DALTON and BYWATER's estimates varied from a negative value to 17%, that of FRAHM and BROWN was also 17%. The estimate of heritability from these data is 0.28 and while still high compared to other work, both the estimate from the component analysis and the realised heritability are very similar. However, there is the problem that any apparent progress did not occur until generations 19 and 20. The regression of TWW on generation number was calculated for the control line to see if there was an environmental trend, and the estimate was 0.0485. If this is used as an estimate of environmental effects, it accounts for only about 20% of the response.

The analysis of variance provides separate estimates of heritabilities of the selection trait and its components. All show high values (0.18–0.28) which predict potentially good progress from selection. These estimates for numbers born, numbers weaned and average weaning weight (as well as TWW) are all quite high compared to those quoted earlier, although most of these were realised heritabilities as opposed to ANOVA estimates. In addition, the estimates from this experiment are based on full sib information and are almost certainly inflated by a common maternal environment. FALCONER (1963) estimated the heritability of litter size to be 0.08, whereas EISEN (1978) obtained realised heritabilities of 0.16–0.19, although his work was based on litters standardised for rearing size. NAGAI et al (1978) measured nursing ability and considered it to be reasonably heritable (11–16%). The genetic correlations shown in Table 4 differ from some of the estimates reported in other work. EISEN (1978) reported a realised genetic correlation between individual three week and litter size of 0.23, in agreement with the range of 0.24–0.44 estimated by JOAKIMSEN and BAKER (1977). The estimate in this analysis between N and W is negative (−0.333) and is that between W and number born (−0.209). However the standard errors attached to these estimates are sufficiently high that these need not be considered different from zero.

It would be reasonable to conclude that the experiment failed to show significant response to selection for TWW, since changes in the last two generations out of 20 can hardly be considered to be acceptable as a response. There are possible reasons for this lack of response and these are discussed in detail by EISEN (1981). Firstly the mating structure used in this experiment is shown by EISEN to be less efficient than selection with families of half sisters, which in turn is less efficient than selection on individual performance. Secondly, EISEN points out that a negative correlation between litter age and individual weight can result in a substantial reduction in response.

However, considered on a generation basis (see Figure 4) the numbers weaned and average individual weaning weight appeared to change synchronously rather than the opposite, whereas it might have been expected that without standardisation of litter size there might have been a seeaswing effect from maternal influences. It is possible that the practice of selecting the haviest females from within the selected litter overcame the potential reciprocity, but is has not been possible to examine this in detail.

The design of the experiment attempted to minimise inbreeding but numerical deficiences could cause some additional effects over the 20 generations. The calculated effective number was 55.8 (compared to the expected 60) which gives a rate of inbreeding of 0.009 per generation and 16.5% over the experiment as a whole. This level is not high by laboratory animals standard, and it would not be expected to have had much effect. The control line had a similar Ne of 52.9. The rate of inbreeding is very similar to the selected line, and the drift variance was 0.0056.

It is difficult to assess the usefulness of this experiment in indicating potential in other species, particularly farm livestock such as pigs or sheep. The point at which mouse

experimental work may not be directly relevant to pigs is because mice were used (and this experiment is no exception) at ages which may not be equivalent to the age of interest in pigs. For example, though weaning in both species is done at a biologically sensible age, they may not be metabolically equivalent.

The trait as measured is effectively a trait of the dam, as indicated by the lack of any sire effect in the analysis of variance. There is therefore no adjustment required due to differences due to sire of litter, and if this were confirmed in pigs, selection would be made easier. The possibility of altering responses by different weightings in the index requires further study. We have inadequate information on correlated responses due to selection for TWW. The most desirable trait for improvement in terms of sow productivity is not clear (BICHARD and COATES 1979) but if not total weaning weight, then a trait of which it is a component. The evidence from mouse experiments when selecting for reproductive performance is that selecting in pigs, using an adequate selection criterion, offers some hope of reasonable progress.

The evidence of the experiment supports some of EISEN's (1981) conclusions. Selection for TWW is expected to yield some response, but selection should be more efficient if based on an appropriate index of the component traits. EISEN concludes that indirect selection for TWW, using litter size at birth and standardising litter size for rearing, may be more effective than direct selection, because it avoids the consequences of negative correlations between litter size and individual body weight. The modest response (if any) in the experiment reported here indicates that alternative approaches should be considered.

Summary

Four lines of mice selected to their limit for 6 week body weight were crossed to form the base population for further selection for the total weaning weight (TWW) of the litter at 3 weks of age. Selection on a within-family basis was carried out for 20 generations, each generation comprising fifteen families each of four full-sib dams. The four females of each family were mated to two full-sib males, avoiding close relatives. A control line was available from generation 5 onwards.

Analysis of generation means and deviations from the control showed that little real progress had been made, and what progress there was appeared only in the last two generations. Despite this, estimates of realised heritabilities from this analysis ranged from 23–58%, with the estimate from the control line being 28%.

Analysis of variance within generations showed that the effect of the litter's sire on TWW was negligible. Heritability of TWW from this analysis was 0.25 ± 0.09 and for numbers born, numbers weaned (N) and average individual weaning weight (W), 0.28 ± 0.09, 0.18 ± 0.09 and 0.28 ± 0.09 respectively. Both N and W were favourably correlated with TWW, but unfavourably with each other, both genetically and phenotypically.

The trait as measured is effectively a trait of the dam and if this were confirmed in farm livestock, selection would be made easier. But the poor response to selection in this experiment supports some of the conclusions of EISEN (1981) that alternative approaches should be considered, particularly for farm livestock.

Résumé

Sélection sur poids du sevrage des souris et son importance pour animaux domestiques

Quatre lignées de souris avec plateau de sélection pour poids du corps de 6 semaines ont été croisées pour former la population de base pour sélection ultérieure pour le poids total au sevrage à l'âge de 3 semaines. (TWW) La sélection entre les familles fut entreprise pour 20 générations, chaque génération comprenant 15 familles, chaque famille 4 pleine-soeurs. Les quatre femelles de chaque famille ont été accouplées à deux plein-frères, ainsi fut évitée une parenté étroite. Une lignée de contrôle fut disponible à partir de la génération 5.

L'analyse des moyennes des générations et des déviations du contrôle ne fournissait que peu de références pour un progrès effectif et ce progrès minime n'apparaissait que dans les deux dernières générations. Malgré cela, les estimations des héritabilités réalisées de cette analyse rangeaient de 23 à 58%, l'estimation de la lignée de contrôle était de 28%.

L'analyse de variance en-dedans des générations montrait que l'effet du père sur TWW était négligeable. L'héritabilité de TWW de cette analyse était 0,25 ± 0,09 et pour les nombres né (N) et poids au sevrage moyen individuel (W), 0,28 ± 0,09, 0,18 ± 0,09 et 0,28 ± 0,09. N aussi bien que W était

favorablement corrélé avec TWW, mais défavorablement avec chaque autre et certes aussi bien génétiquement que phénotypiquement.

Le caractère examiné est effectivement un caractère de l'animal-mère et si cela était confirmé pour animaux domestiques, la sélection serait nettement plus facile à exécuter. La réaction insuffisante sur la sélection dans cette expérience soutient quelques conclusions de Eisen (1981), que les approches alternatifs doivent être considérés, notamment pour animaux domestiques.

Resumen

Selección para peso total al destete en ratones y su implicación para animales domésticos

Cuatro lineas de ratones seleccionados para peso del cuerpo a las seis semanas de edad fueron cruzadas con el fin de obtener la población base para una selección ulterior a peso total de la camada al destete (TWW) a las tres semanas de edad. La selección entre familias fué realizada durante 20 generaciones, comprendiendo cada generación 15 familias y cada familia cuatro hermanas enteras. Las cuatro hembras de cada familia fueron apareadas con dos machos hermanos enteros tratandose de evitar un parentesco estrecho. Una linea testigo estaba disponible a partir de la quinta generación.

Un análisis de los promedios de generaciones y desviaciones del testigo demostró que se habia hecho poco progreso. Donde hubo un progreso este aparecía solamente en las dos últimas generaciones. No obstante, las estimaciones de heredabilidades realizadas en este análisis alcanzaron de 23–58%, siendo de 28% para la linea testigo.

El análisis de varianzas entre generaciones mostró que el efecto del macho semental sobre TWW era negligible. La heredabilidad para TWW era 0.25 ± 0.09, y para número de animales nacidos, número de destetados (N) así como peso medio al destete (W) era 0.28 ± 0.09, 0.18 ± 0.09 y 0.28 ± 0.09 respectivamente. N y W estaban favorablemente correlacionados con TWW, pero desfavorablemente entre sí, tanto genética como fenotipicamente.

La característica investigada efectivamente es una que debe ser atribuida a la hembra-madre. Si esto pudiese ser confirmado para animales domésticos, la selección sería mucho más fácil. Con todo, la pobre respuesta a la selección en este experimento apoya algunas de las conclusiones de Eisen (1981), o sea que deben ser considerados otros aspectos, especialmente tratándose de animales domésticos.

Zusammenfassung

Selektion auf Wurfabsatzgewicht bei Mäusen und ihre Bedeutung für Haustiere

4 Mäuselinien mit Selektionsplateau für 6-Wochen-Gewicht wurden gekreuzt und dienten als Basispopulation für weitere Selektion auf 3-Wochen-Wurfabsatzgewicht (TWW). Die Selektion zwischen Wurfgeschwistern wurde durch 20 Generationen vorgenommen, wobei jede Generation 15 Familien aus je 4 Vollschwestern umfaßte. Die 4 Vollschwestern wurden an 2 Vollbrüder gepaart, wobei enge Verwandtschaft vermieden worden ist. Eine Kontrollinie war von der Generation 5 an verfügbar.

Die Analyse der Generationsmittelwerte und Abweichungen von der Kontrolle ergaben kaum Hinweise auf tatsächlichen Fortschritt und der geringe erzielte Fortschritt schien nur in den letzten 2 Generationen entstanden zu sein. Trotzdem wurden realisierte Heritabilitätswerte in der Größenordnung von 23 bis 58% gefunden, wobei die Schätzung der Kontrollinie 28% war.

Die Varianzanalyse innerhalb Generationen zeigte, daß der Wufvater einen zu vernachlässigenden Einfluß auf TWW hatte. Die Heritabilität von TWW aus dieser Analyse war 0,25 ± 0,09 und für die Zahl geboren, Zahl abgesetzt (N) und durchschnittliches individuelles Absatzgewicht (W), 0,28 ± 0,09, 0,18 ± 0,09 und 0,28 ± 0,09. Sowohl N wie auch W waren mit TWW in günstigem Sinn korreliert, aber ungünstig miteinander, und zwar sowohl auf genetischer wie auch auf phänotypischer Basis.

Die erhobene Eigenschaft ist tatsächlich ein Merkmal des Muttertieres und wenn dies auch für Haustiere gilt, könnte Selektion wesentlich leichter durchgeführt werden. Die mangelhafte Reaktion auf die Selektion in diesem Versuch unterstützt einige Schlußfolgerungen von Eisen (1981) dahingehend, daß alternative Ansätze berücksichtigt werden sollten, besonders bei Haustieren.

References

Barber, R. S.; Braude, R.; Mitchell, K. G., 1955: Studies on milk production of Large White pigs. J. Agric. Sci. **46**, 97–118.

Bichard, M.; Coates, A., 1979: The theory of creation of prolific lines. EAAP/MLC/ABRO European Pig Testing Conference, Paper E5.2.

Dalton, D. G.; Bywater, T. L., 1963: The effect of selection for litter size and litter weight at weaning in mice maintained on two diets. Anim. Prod. **5**, 317–326.

Eisen, E. J., 1973: Direct and maternal genetic responses resulting from selection for 12-day litter weight in mice. Can. J. Genet. Cytol. **15**, 483–490.

EISEN, E. J., 1974: A laboratory mouse as a mammalian model for the genetics of growth. Proc. 1st Wld. Cong. on Gen. applied to Live Prod., Vol. **1**, 467–492.
EISEN, E. J., 1978: Single-trait and antagonistic index selection for litter size and body weight in mice. Genetics **88**, 781–811.
EISEN, E. J., 1981: Predicting selection response for total litter weight. Z. Tierzüchtg. Züchtgsbiol. **98**, 55–76.
EISEN, E. J.; LEAGATES, J. E.; ROBISON, O. W., 1970: Selection for 12-day litter weight in mice. Genetics **64**, 511–532.
FALCONER, D. S., 1955: Patterns of response in selection experiments with mice. Cold Spring Harbour Symp. **20**, 178–196.
FALCONER, D. S., 1960a: The genetics of litter size in mice. J. Cell. Compar. Physiol. **56**, Suppl. 1, 153–167.
FALCONER, D. S., 1963: Qualitatively different responses to selection in opposite directions. Statistical Genetics and Plant Breeding NAS-NRC 982.
FRAHM, R. R.; BROWN, M. A., 1975: Selection for increased preweaning and postweaning weight gain in mice. J. Anim Sci. **41**, 33–42.
JOAKIMSEN, O.; BAKER, R. L., 1977: Selection for litter size in mice. Acta. Agric. Scand. **27**, 301–318.
LEGATES, J. E.; FARTHING, B. R., 1972: Selection for growth and maternal performance in mice. J. Anim. Sci. **21**, Abstr. 974.
NAGAI, J.; EISEN, E. J.; EMSLEY, J. A. B.; MCALLISTER, A. J., 1978: Selection for nursing ability and adult weight in mice. Genetics **88**, 761–780.
ROBERTS, R. C., 1965: Some contributions of the laboratory mouse to animal breeding research. Part 1 and Part 2. A. B. A. **33**, 339–353 and 515–526.
ROBERTS, R. C., 1967: The limits to artificial selection for body weight in the mouse. III. Selection from crosses between previously selected lines. Genet. Res., Camb. **9**, 73–85.
ROBINSON, W. A.; WHITE, J. M.; VINSON, W. E., 1972: Response to selection for 12-day litter weight in mice. J. Anim. Sci. **35**, 182, Abst. 66.

Authors' addresses: D. E. STEANE, Meat and Livestock Commission, PO Box 44, Queensway House, Bletchley MK2 2 EF, England. R. C. ROBERTS, ARC Animal Breeding Research Organisation, Kings Buildings, West Mains Road, Edinburgh EH 9 3JQ, Scotland.

PAPER 14.

The lifetime growth and reproduction of selected strains of mice.

Hered. 16, 369-381. 1961.

by

R.C. ROBERTS

The Lifetime Growth and Reproduction of Selected Strains of Mice

R. C. ROBERTS

Agricultural Research Council Unit of Animal Genetics, Institute of Animal Genetics, Edinburgh, 9

Reprinted from HEREDITY, Volume 16 Part 4 pp. 369-381 November 1961

THE LIFETIME GROWTH AND REPRODUCTION OF SELECTED STRAINS OF MICE

R. C. ROBERTS
Agricultural Research Council Unit of Animal Genetics, Institute of Animal Genetics, Edinburgh, 9

Received 30.iii.61

1. INTRODUCTION

Most experiments in quantitative genetics perforce emphasise progress per unit of time, with the resultant quick turnover of generations. Thus, for instance, mouse selection experiments are commonly based on the use of first litters only, where nothing is known of the potential lifetime performance of the animals. Indeed, such information is probably irrelevant from the restricted point of view of a particular study. Yet, it is important to know how experimental procedures with the concomitant changes in genic arrays affect the natural fitness of the organism. In so far as such procedures may disturb gene frequencies in a population at equilibrium, fitness may be expected to decline, on the model of genetic homeostasis proposed by Lerner (1954). It is possible therefore that selection for any character in either direction may lead to a decline in natural fitness. Natural fitness of course has many components whose individual identities must often be obscure. Cumulatively, however, they are perhaps most clearly related to the total reproductive capacity of the organism, measured over the organism's lifetime.

The present study, which is essentially descriptive in nature, comprises a limited examination of these questions. Samples of mouse strains selected for high and low weight were placed aside and allowed to complete their reproductive life. Two main points were examined:

1. The effect of selection for six-week weight on subsequent growth, which may determine or at least affect reproductive capacity.
2. The effect of selection for weight on various aspects of longevity, and particularly on the total number of progeny weaned.

2. MATERIAL

Six groups of mice were employed in a comparative study. The designation and derivation of the groups were as follows.

1. RCL, a cross between Goodale's and MacArthur's large mice, and selected further for high six-week weight for 10 generations. The origin of the stock is described more fully by Falconer and King (1953).
2. MS, MacArthur's small mice, selected further in this laboratory for low six-week weight for 17 generations.
3. MXR, the F_1 generation of a cross between MS and RCL; reciprocal crosses are represented equally in this group.

4. NF, selected for high six-week weight for 27 generations.
5. NS, selected for low six-week weight for 22 generations.
6. NC, an unselected control stock from the same foundation as NF and NS, kept for 12 generations when this study started.

The last three stocks are described in greater detail by Falconer (1953).

Ten pair matings of each kind were set up and records of weight and litter production were kept. When an animal died, a replacement was provided in order to obtain records from the surviving member of the pair. Death usually resulted from natural causes, although animals were killed if they were in obvious distress and death in any case seemed imminent.

Much of the material, by nature of its erratic reduction until only one animal remained, does not lend itself easily to statistical treatment. Attempts were made to overcome some of the difficulties by the use of various transformations, invariably without much success. Fortunately, however, the main conclusions are often self-evident from the raw data, the presentation of which alone then suffices.

3. RESULTS

(i) *Length of life*

The age at death in days was calculated for all animals. Five mice whose deaths were due to accidental causes have been excluded. The results are shown in table 1, with the appropriate analysis of variance in table 2.

TABLE 1

Mean age at death in days

Stock	Males	Females
RCL	474·2±37·2	283·8±29·7
MS	700·2±44·1	452·2±52·8
MXR	683·1±66·1	471·4±60·5
NF	759·3±62·9	730·6±75·2
NS	900·3±99·4	747·3±73·0
NC	492·7±78·9	545·1±80·9

TABLE 2

Analysis of variance of mean age at death

	df	SS	MS	*p*
Total	114	7 190 069	...	...
Between sexes	1	414 812	414 812	<0.01
Between stocks	5	2 277 620	455 524	<0.001
Interaction	5	396 135	79 227	$>0.05<0.10$
Error	103	4 101 502	39 820	

Among the stocks tested in this trial, there was significant variation between stocks around a mean lifespan of 603 days. Furthermore, under the conditions prevailing in mating cages, males were significantly longer-lived than females, the weighted difference being 128 days. There is some suggestion that this difference is not constant

from stock to stock, though the interaction mean square is but barely significant at the 10 per cent. level.

Two comparisons enable us to examine the effect on length of life of selection for body weight, that between RCL and MS and that between NF and NS. In both cases, small mice were longer-lived though not significantly so in the latter case. Unexpectedly, among the N stocks selection in either direction increased the lifespan. In the MXR stock, the character displayed considerable heterosis, which is perhaps to be expected. The crossbreds equal almost exactly the longer-lived parental stock. The longest-lived animal in the whole experiment was an NS male, who lived to 1330 days (3 years 8 months). Four animals of the NS stock, two males and two females, exceeded 1000 days.

(ii) *Patterns of growth*

All animals were weighed when weaned at three weeks, and then at six weeks which is the usual age at mating. Males of the N stocks were weighed again at nine weeks, and males of all stocks at twelve weeks. Thereafter all males were weighed at four-week intervals until death. After six weeks of age, females were weighed immediately the birth of a litter was recorded, to avoid the obvious variation due to pregnancy. This, however, meant that many females lived over long periods, especially towards the end of their life, without a weight being recorded.

As animals died, the mean weight for each stock became determined by successively smaller numbers as time progressed, until ultimately only one animal remained. For this reason, the results are presented in fig. 1 as the growth curves of individual male mice. The distribution of growth patterns between stocks is sufficiently distinct, compared to variation within a stock, for the main conclusions to be drawn without any statistical refinements. However, to compensate for any bias in the apparent trends due to a possible correlation between weight and life-span within a stock, cumulative growth curves for males of each stock are shown in fig. 2. These curves were drawn by accumulating the growth of survivors over successive time intervals and are therefore largely independent of the absolute weight of the survivors at the time. The curves in fig. 2 are discontinued when they become determined by fewer than three mice.

We shall not consider the effect of the selection on six-week weight, which is another topic, but rather the consequences of the selection on further growth. Considering initially the weights of males, it is seen that growth was not complete at six weeks, and that the differences between large and small stocks increased rather than diminished, at least for a while. It is at once apparent that selection on total growth up to six weeks has resulted in vastly different mature sizes. Further, inspection of the growth curves shows that even within a stock, there

existed a correlation between six-week weight and maximum weight. Yet, this is not the whole story. A comparison of the two large stocks (RCL and NF) shows that both attained the same mature weight

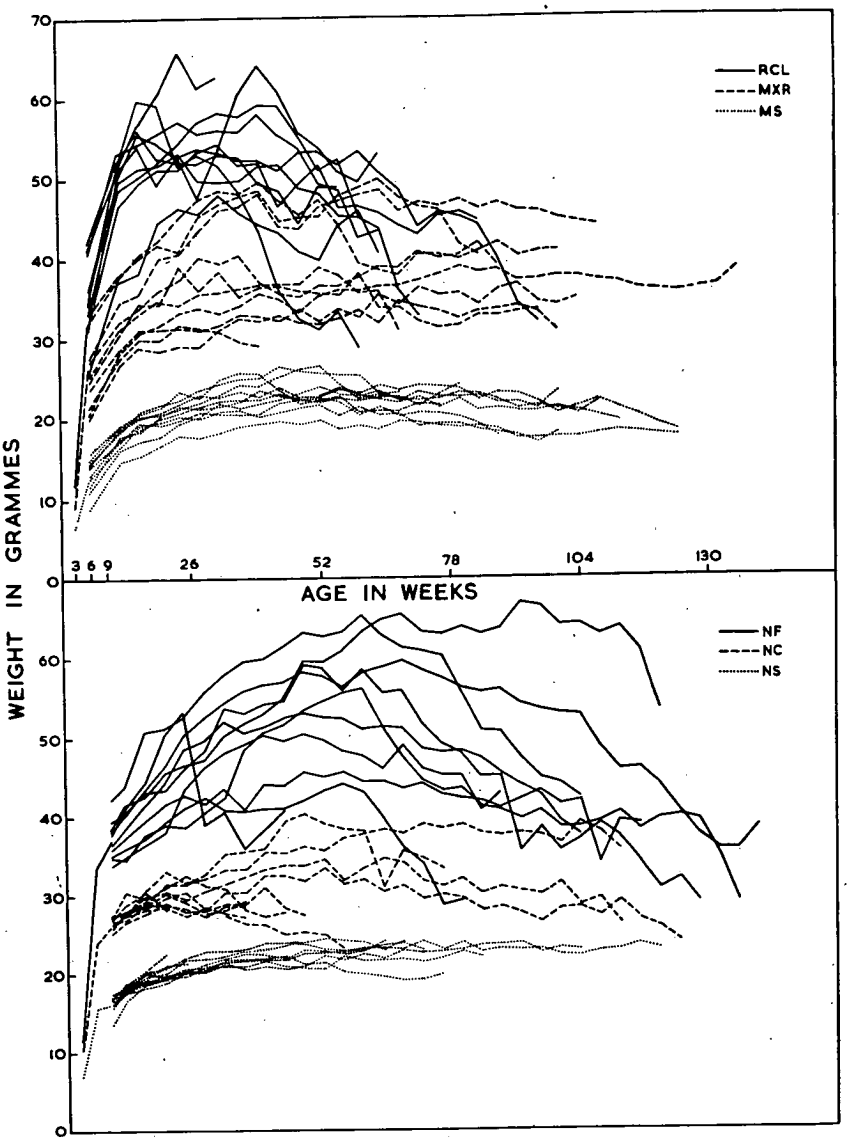


FIG. 1.—Individual growth curves of male mice, showing differences in growth pattern between stocks. (Five NS males lived longer than shown, but some of their weights inadvertently were not recorded.)

but at different ages, in the former at approximately six months of age, and in the latter at approximately one year. This is at least presumptive evidence that mature weight and the path whereby it is reached are to some degree under separate genetic control. On this

point, it can be noted that the RCL males reached their maximum weight at a much younger age than males of other stocks.

Once the maximum weight had been attained, a conspicuous difference in pattern emerged between the two heavy stocks and the others; the former, almost immediately and without exception, began to show a decline in weight which continued without arrest until death. The two light stocks and also the unselected and crossbred stocks, in contrast, showed no decline in weight and retained their

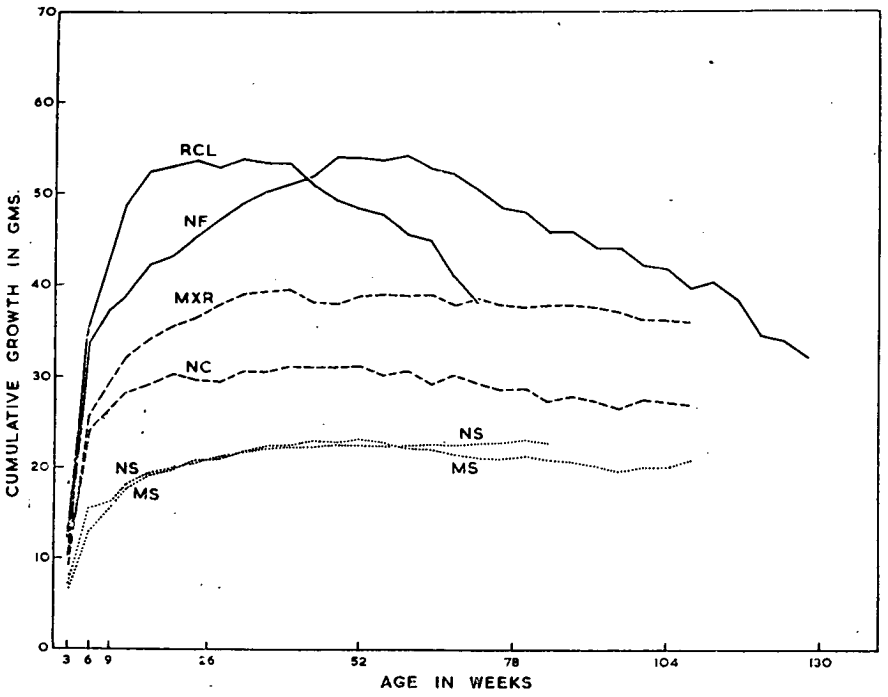


FIG. 2.—Cumulative growth curves of male mice for each stock (explanation in text).

maximum weight almost until death. It seems probable that this difference in weight pattern is related to the difference in fatness; the large mice contain much larger stores of fat, relative to body weight, than small mice (Fowler, 1958). The loss of weight of the large stocks in later life was probably due to the depletion of accumulated fat.

The crossbred stock (MXR) showed no evidence of heterosis with respect to mature weight. It did, however, maintain this weight, and it eventually exceeded the level of the heavier parental strain. Whether weight can be regarded as a heterotic character depends therefore on the age at which the character is measured; alternatively, one may argue that the nature of the " character "—in this case, weight—may alter, with age. For instance, it is quite possible that the MXR stock showed heterosis with respect to bodily dimensions at all ages, but failed to accumulate as much fat as the RCL stock.

The MXR group consisted of an equal number of the two reciprocal crosses. Whereas the group, as a group, was quite distinct from either of the two parental stocks, there was some evidence that within the group mice derived from RCL mothers were heavier than those derived from MS mothers. At maximum weight, the average difference was of the order of five grammes, and the two reciprocal types converged only trivially as they aged, indicating a permanent maternal effect on weight. Though the numbers are quite insufficient to establish this maternal effect as being statistically significant, it is proportionately of the same order of magnitude as that found by Brumby (1960) in analogous groups of mice at twelve weeks of age.

It was explained earlier that the weights of females were somewhat more ambiguously recorded than those of males. In general, however, they provided confirmatory evidence for the conclusions drawn above. All stocks showed that at six weeks, males were appreciably heavier than females. In the two light and two intermediate stocks, however, the ranking of the sexes almost immediately became reversed. This could easily be attributed to a peculiarity of the post-partum weight recorded in females, were it not for the fact that the two heavy stocks did not exhibit the same phenomenon. A possible explanation might be that breeding females of the RCL and NF stocks did not accumulate the same amount of fat as males of those stocks.

(iii) *Reproductive performance*

The data referring to the reproductive history of the stocks are summarised in tables 3 and 4, which refer to females only. The first point, not revealed in the tables, is that stocks did not differ significantly in the age at which the first litter was born. This would almost certainly not have been the case had not mating been delayed for all stocks until they were on average about eight weeks old. Having thus begun on an equal footing, four major factors govern the reproductive capacity of the stocks—the length of breeding life, the interval between litters, the average size of the litters at birth and the proportion of those born that survive to be weaned.

It is seen that the female mice in this trial stopped breeding at an average age of 300 days or so, but the variation among stocks is highly significant. Reproduction ceased at a younger age in the large stocks than in the small stocks. Further, referring back to table 1, some stocks, notably the NF, lived for a considerable time after reproduction had ceased. The correlation between the age at last litter and female longevity, with respect to stock means, is less than 0·5, despite an obvious causal relationship. However, this estimate, based on only six stocks, cannot be very accurate. Neither of the large stocks produced litters beyond the point when they would be expected to show the decline in weight characteristic of males of those stocks. Unfortunately, for reasons explained earlier, no weights

were obtained from females in later life to determine whether in fact they did lose weight.

The conclusions with respect to the age at which the last litter was born are clear. Firstly, selection for large size decreased the length of reproductive life while selection for small size increased it.

TABLE 3

Lifetime reproduction of females—numbers born alive

Stock	Mean age (in days) at last litter	Mean no. of litters	Mean interval (in days) between litters *	Mean total no. born	Mean litter size at birth †
RCL	201·2±30·9	3·4±0·92	42·65	24·2± 6·77	7·1
MS	295·8±51·2	10·6±1·49	22·74	49·0± 7·37	4·6
MXR	377·6±43·3	11·0±1·52	29·36	102·5±10·35	9·3
NF	238·4±28·9	5·4±0·83	34·44	33·0± 5·77	6·1
NS	428·0±44·7	11·9±1·47	30·92	48·8± 6·65	4·1
NC	307·2±27·1	8·0±0·54	32·00	44·8± 5·57	5·6

* $\frac{\text{Mean age at last litter—Mean age at mating}}{\text{Mean no. of litters}}$

† $\frac{\text{Mean total no. born}}{\text{Mean no. of litters}}$

TABLE 4

Lifetime reproduction of females—numbers and weight of offspring weaned

Stock	Mean total no. weaned	Proportion weaned	Mean total wt. (in gm.) of offspring weaned	Mean total wt. (in gm.) weaned by 183 days	Proportion of total wt. weaned by 183 days
RCL	18·6±5·70	0·77±0·027	198·99± 56·13	159·14±32·12	0·80
MS	31·1±6·30	0·63±0·022	205·70± 41·85	100·53±16·31	0·49
MXR	93·4±9·27	0·91±0·009	847·24±106·14	414·68±36·87	0·49
NF	18·8±4·57	0·57±0·027	176·54± 36·57	146·95±26·39	0·83
NS	37·5±7·98	0·77±0·018	265·45± 53·89	113·25±16·68	0·43
NC	33·4±5·16	0·75±0·020	298·31± 48·75	198·97±32·99	0·67

Secondly, reproductive longevity, as illustrated by the MXR stock, displays striking heterosis.

The length of reproductive life is reflected, to a large extent in the mean number of litters born to females of particular stocks. It is seen that the large stocks do not compensate for their shorter reproductive life by a more rapid litter production. On the contrary, the disparity between the large and the small stocks is magnified rather than diminished. However, the large stocks make up some leeway by producing larger litters, few though they may be. But the advantage of the small stocks in the number of litters born is a telling

one, and their superiority over the large stocks in the total number of offspring born is in no doubt.

The fitness of an animal depends of course not so much on the number of progeny to which it gives birth as on the number of those progeny that reach sexual maturity and themselves reproduce. However, it is common experience that relatively few losses occur in the laboratory mouse after weaning, and that sterility also is comparatively rare. The total number of offspring weaned therefore provides some assessment, though by no means an exact one, of the reproductive fitness of a mouse. When the appropriate column in table 4 is examined, the small stocks are seen to have retained their advantage in the total number of offspring by weaning time. It is further seen that this is mostly due to the initial advantage in number born, and that the variation in the proportion of offspring that survive to weaning is not correlated with large and small size. Perhaps the most impressive feature of the data is the striking heterosis displayed by the MXR stock, which weaned three times as many young as the better of its two parental stocks. The crossbred animals were very successful in all aspects of reproduction, and outstandingly so in the mean litter size at birth and in rearing those litters to weaning.

The pattern of reproduction in the mouse is well-known. The first litter on average is submaximal owing to fewer ova being shed. Litter size then remains at a fairly stable maximal level over three parities or more, but eventually it gradually declines. The decline is reported to be due at least in part to an increased incidence of foetal mortality (Hollander and Strong, 1950; Wanke, 1939; Murray, 1934). The mean litter sizes at birth of successive parities for each stock are shown in fig. 3. The graphs, which are discontinued when they become determined by fewer than four mice, all conform to the expected pattern. There is some variation in the parity at which the maximum litter size is achieved which, however, does not seem to be correlated with body size. The crossbred stock (MXR), despite retaining its superiority in mean number born over twelve litters, does ultimately fall to the level of the small stocks; in fact, its decline with advancing age is the most marked. Nevertheless, it appears that the onset of senescence, as judged by the decline in litter size with age, is somewhat delayed in the crossbred stock, a finding which fails to support a suggestion by Chai (1959) that hybridisation confers no such effect.

The general conclusion with regard to the reproduction of the stocks tested in this trial is that mice selected for small size have a far higher reproductive rate than mice selected for large size, and possibly a higher rate than the unselected control mice. It has further been shown that the advantage of small mice rests entirely on the number of litters that they produce, which in turn can be related to the length of reproductive life. The one crossbred stock tested displayed the expected heterosis in reproductive capacity.

(iv) *Total weight of offspring weaned*

Those concerned with the applied aspects of this study may legitimately ask whether the number of offspring weaned is as important as the total weight of those offspring, especially the total weight within a given time from mating. This is primarily an economic question for which there is no general answer, but the data collected in this study are presented in table 4. It is seen that the advantage

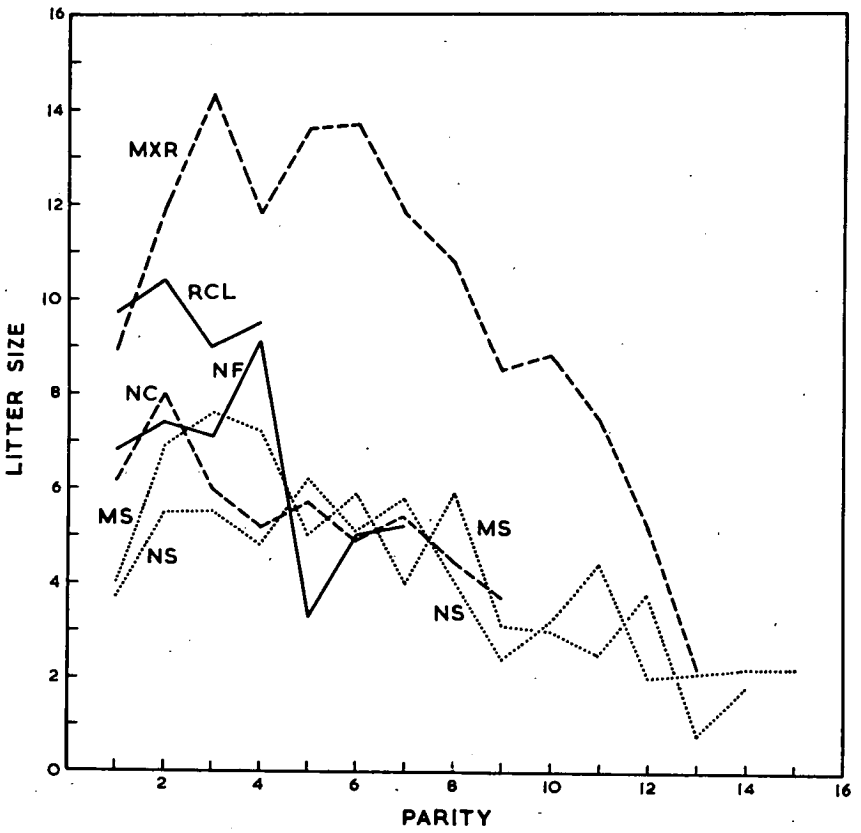


FIG. 3.—The relationship between litter size at birth and parity, for each stock.

of the small stocks in the total number weaned is sufficient to give them a superiority also in the total weight of these offspring at weaning. However, in neither of the two comparisons between large and small stocks is the difference statistically significant; in fact, the superiority of the MS stock over the RCL is quite trivial. Furthermore, as shown previously, the small stocks took a considerably longer time to accumulate those weights. For this reason, the cumulative weaning weights up to the age of 183 days (6 months) were examined. This period was chosen because by then all stocks are past their peak reproductive capacity, and are therefore at a stage when matings would be terminated under normal mouse management. The

appropriate column in table 4 shows the large stocks in a more favourable light. At this age, they exceed the performance of the small stocks, though not significantly. The conclusion is that if interest rests on the weight of offspring per female parent, stocks selected for high body weight are probably preferable to stocks selected for small size, which ultimately give more litters but these are smaller and contain lighter mice. A firmer conclusion would invoke economic considerations which are beyond the scope of this paper.

Two final points from table 4 should be mentioned. Firstly, the performance of the crossbred stock (MXR) is again outstandingly high. Secondly, among the N stocks, selection for six-week weight in either direction has resulted in a decrease of the total weight of offspring weaned, compared to the unselected control stock. The immediate reason for this is not hard to find. In the small stock the decrease is directly attributable to a reduction in weight as a result of the selection, while the decrease in the large stock results from the marked reduction in the number of offspring weaned.

4. DISCUSSION

None of the characters described in this paper had been the subject of previous artificial selection in these stocks of mice. The differences found between large and small stocks can therefore be classified as correlated responses to selection for six-week weight. A detailed discussion of the theory of correlated responses would be inappropriate here; a recent exposition is given by Falconer (1960). Briefly, however, a correlated response to selection cannot arise in the absence of a genetic correlation between the two characters, though other factors will also contribute to the magnitude of the response. A genetic correlation implies either that some genes affect both characters, *i.e.* are pleiotropic, or that genes affecting the characters separately are linked predominantly in the one phase. Whereas in the selected lines, it might be expected that equilibrium would have been reached between the coupling and repulsion phases, it might well be that genetic correlations as a result of linkage would be important in the crossbred stock examined in this study.

Selection on total growth up to six weeks has been shown in the previous section to have affected a number of other characters as well. Not surprisingly, perhaps, it has had a marked effect on subsequent growth and has led to different maximum weights, though the proportional differences between the large and small stocks are substantially the same at six weeks and at mature weight. From this alone, it would not be unreasonable to argue that the same genetic system controls both six-week and mature weights. But the different growth pattern of the two large stocks clearly suggests at least some genetic independence between growth rate and mature weight, and

thus shows broad agreement with the hypothesis of mammalian growth proposed by Dickinson (1960).

Selection for body weight had a marked effect on length of life and particularly on the length of reproductive life. The reason for this is obscure, but a possible parallel may be drawn from nutritional studies by McCay (1947), Ball *et al.* (1947) and Visscher *et al.* (1952). Briefly, these studies cumulatively show from work with rodents that the restriction of calorie intake through the diet lengthens life and also delays the age of reproductive failure, as judged by the capacity of the females to produce litters. Now, the parallel is pertinent only if the small stocks in this study can be regarded as a biological means of restricting calorie intake. For instance, small mice may utilise a relatively greater portion of their calorie intake to maintain body temperature. If so, then the direct cause of the differences found in length of reproductive life becomes a mechanical one resulting from differences in body size, though the complexities at the physiological level remain.

The finding concerning the relative longevity of large and small stocks is not in accord with that of Chai (1959), whose large strain significantly outlived his small strain. His material, however, was inbred, and is therefore not directly comparable with the present study. Chai found too that a hybrid stock derived from a cross between his large and small strains had a lifespan in excess of either parental strain.

The correlated responses to selection found in this trial were usually in opposite directions in the large and small stocks. On a homeostatic model, discussed earlier, this is not the expectation for characters related to fitness. The general finding suggests therefore that in the material examined here, genetic homeostasis was not a predominating feature of the base populations. There were only two instances in which the correlated responses assumed the same direction in both large and small mice, the comparisons being of course limited to the N stocks which contained a control. Firstly, the mean age at death increased in both selected stocks. The reason for this is obscure. It should, however, be noted that the age at which reproduction ceased, a more direct component of fitness, responded differently in large and small mice. Secondly, the total weight of offspring weaned declined in both selected lines for reasons given earlier. While this character is almost certainly related to some aspects of fitness, it is less obviously so than the total number weaned, where the correlated responses were in opposite directions.

The final conclusion from this study is that selection for a rapid early growth had an adverse effect on reproductive fitness, as judged by the total number of offspring weaned over a lifetime. This reduction in number of offspring resulted mainly from a drastic shortening of the length of reproductive life. Given time, mice selected in the opposite direction showed no reduction in number weaned, though of

course the total weight of offspring was reduced. In terms of practical application to domestic livestock, a good growth rate and high reproductive capacity are frequently twin objectives. The general conclusion is therefore discouraging. However, depending upon economic and biological considerations, the consequences of a shortened reproductive life in some species may not be serious. Further, the apparent negative correlation between growth rate and reproductive capacity may not be so great that its effect could not be overcome or at least diminished by appropriate selection techniques. And lastly, this study confirms that the crossing of suitably selected strains may offer a solution to what might otherwise become stagnating fertility problems.

5. SUMMARY

1. This paper reports the effect of selection for six-week weight in mice on their subsequent growth and on various aspects of their longevity, particularly on the length of reproductive life.

2. The material consisted of ten pair matings each of two large strains, two small strains, a large × small cross and an unselected control strain.

3. The average length of life over all stocks was 1 year 8 months. The mean life span of the small strains exceeded that of the large strains by approximately 6 months, while the crossbred stock equalled almost exactly the better parental strain.

4. The difference in weight between the large and small strains magnified with age though the proportionate difference remained fairly stable. However, the two large strains, once their maximum weight had been achieved, showed a decline in weight which continued until ultimately they fell to the level of the intermediate strains. The cause of the decline, which was not a feature of the other four groups, was probably the depletion of fat reserves.

5. The large strains had a short reproductive life, producing on average only $4\frac{1}{2}$ litters, against 11 or so in the small strains. On account of this the small strains eventually weaned almost twice as many offspring as the large strains.

6. Perhaps the most striking feature of the data was the heterosis displayed by the crossbred stock with respect to reproductive capacity. Compared to the better parental strain, the crossbred stock weaned three times as many offspring whose total weaning weight was four times as great.

7. The findings are discussed in relation to the original selection for high and low six-week weight.

Acknowledgments.—I am greatly indebted to Dr D. S. Falconer for making available records which made this study possible; he collected the records from mice providing for other requirements. I wish to thank him also for his valuable comments on the manuscript.

I wish to thank Mr E. D. Roberts for drawing the figures; I appreciate his patience with the cumbersome fig. 1.

6. REFERENCES

BALL, Z. B., BARNES, R. H., AND VISSCHER, M. B. 1947. The effects of dietary calorie restriction on maturity and senescence, with particular reference to fertility and longevity. *Amer. J. Physiol., 150,* 511-519.

BRUMBY, P. J. 1960. The influence of the maternal environment on growth in mice. *Heredity, 14,* 1-18.

CHAI, C. K. 1959. Life span in inbred and hybrid mice. *J. Hered., 50,* 203-208.

DICKINSON, A. G. 1960. Some genetic implications of maternal effects—an hypothesis of mammalian growth. *J. Agric. Sci., 54,* 378-390.

FALCONER, D. S. 1953. Selection for large and small size in mice. *J. Genet., 51,* 470-501.

FALCONER, D. S. 1960. *Introduction to Quantitative Genetics.* Oliver and Boyd, Edinburgh.

FALCONER, D. S., AND KING, J. W. B. 1953. A study of selection limits in the mouse. *J. Genet., 51,* 561-581.

FOWLER, R. E. 1958. The growth and carcass composition of strains of mice selected for large and small body size. *J. Agric. Sci., 51,* 137-148.

HOLLANDER, W. F., AND STRONG, L. C. 1950. Intra-uterine mortality and placental fusions in the mouse. *J. exp. Zool., 115,* 131-149.

LERNER, I. M. 1954. *Genetic Homoestasis.* Oliver and Boyd, Edinburgh.

McCAY, C. M. 1947. Effects of restricted feeding upon aging and chronic diseases in rats and dogs. *Amer. J. Publ. Hlth., 37,* 521-528.

MURRAY, W. S. 1934. The breeding behaviour of the dilute brown stock of mice (Little *dba*). *Amer. J. Cancer, 20,* 573-593.

VISSCHER, M. B., KING, J. T., AND LEE, C. P. 1952. Further studies on influence of age and diet upon reproductive senescence in strain A female mice. *Amer. J. Physiol., 170,* 72-76.

WANKE, S. 1939. Einfluss der Wurfreihenfolge auf die Wurfgröse sowie der Wurfgrösse und der Wurfreihenfolge auf die Sterblichkeit von weissen Mäusen während der Saugezeit. *Z. Tierz. ZüchtBiol., 42,* 269-280.

PRINTED IN GREAT BRITAIN BY OLIVER AND BOYD LTD., EDINBURGH

PAPER 15.

Growth regulation in chimaeras between large and small mice.

Nature, Vol. 260, No. 5548, 244-245.
1976

by

R.C. ROBERTS, D.S. FALCONER, PATRICIA BOWMAN
& I.K. GAULD.

(Reprinted from Nature, Vol. 260, No. 5548, pp. 244–245, March 18, 1976)

Growth regulation in chimaeras between large and small mice

AGGREGATION chimaeras between strains of mice differing in body size may be used to determine whether the regulation of growth is determined by the relative proportions of cells deriving from the constituent strains. Should this be so, several questions arise; is growth regulated by the proportions in the body as a whole, or is the cellular composition of particular organs (or tissues) more critical? If critical organs or tissues are established, is the regulation then a systemic property, or does the response of other organs and tissues depend further on their own cellular composition? We describe here preliminary findings which indicate the answers to some of these questions and suggest ways of answering the others.

Aggregation chimaeras were made, by methods described previously[1], between mice from strains selected for large (L) and small (S) body size, and between L and unselected control (C) mice; the history of the strains is given in ref. 2. Some C⟷C chimaeras are included for comparison. Fifty overt chimaeras, all from C-strain foster mothers, were distributed among the three classes in the numbers shown in Table 1. Overt chimaerism was detected by marking the constituent strains with contrasting coat colours, either albino or coloured. Coat colours and body size were used reciprocally in chimaeras between L and S and since there was no difference between the reciprocals, they were pooled. In chimaeras between L and C, the latter was always the albino component. The proportion of albino in the coat was scored as described in ref. 3. We do not include data from successful aggregations which later resulted in single coat colours, but those mice ruled out any general tendency for L cells to outgrow S cells, or vice versa. This is supported by the fact that the mean proportion of L in the coat, in chimaeras both with S and with C, is close to 50% (Table 1).

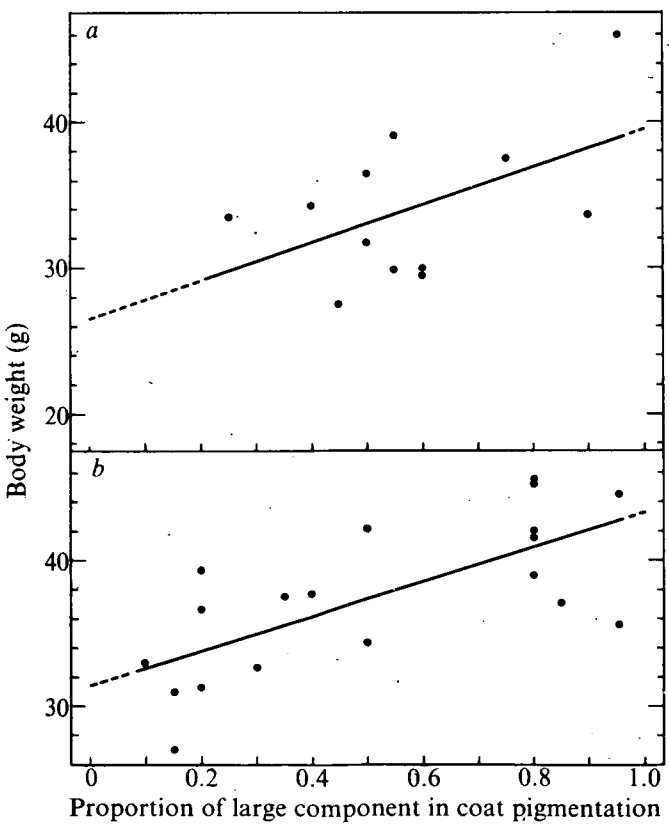


Fig. 1 Relationship between body weight and proportion of L component in coat pigmentation. Values of regression coefficients: in S⟷L (*a*), 13.0±6.9; in C⟷L (*b*), 11.7±2.9.

All data reported here were recorded when the mice were 6 weeks old. All female body weights were adjusted to male equivalents, males being 20% heavier in these strains. The weights of the constituent strains were adjusted further by regression to litter sizes equal to those in which the chimaeras were born, the weights thus seeming greater than those published earlier[2].

Examination of the data revealed a positive regression of body weight on the proportion of the L genotype in the variegation score of the coat (Fig. 1), the corresponding correlations being 0.51 and 0.70 in S⟷L and C⟷L chimaeras, respectively. Cumulatively, there is no doubt about the significance of these regressions, and two conclusions follow. First, body size is linearly and directly proportional to the cellular composition of the coat of the chimaera. Second, since melanocytes themselves can hardly be determinants of growth, the observed correlation must reflect a fundamental correlation between the proportions of melanocytes in the coat and the corresponding proportions of L and S cells in whatever tissue(s) regulate growth.

A critical observation from Table 1 is that S⟷L and C⟷L chimaeras are 2.5 times more variable in weight than C⟷C chimaeras, whose own variance is about the expected normal value. A source of variance in chimaeras between different strains, not present in other mice, is the variation between individuals in the proportions of the two cell types. For coat colour, this proportion varied between 0 and 1, and covered most of that range even if the single colours were excluded. If the same kind of variance in proportion applies to weight-controlling tissue, the inflated variance of weight is readily explicable. It is therefore important to estimate the variance of the proportions of cells in the weight-controlling tissue.

We now define the 'chimaeric genotype' (G) for growth in terms of the proportions of the two types of cells. Let

$$G = PL + (1 - P)S$$

where P is the proportion of large cells (L), leaving $(1 - P)$

Table 1 Body weights and proportion of L component in coats

Constituent lines	L	C	S
Adjusted weights (g)	36.7	27.3	18.3
Chimaera type	S⟷L	C⟷L	C⟷C
n	12	19	19
Coat: proportion L	0.58	0.52	(0.28)*
variance	0.040	0.096	(0.075)*
Weight (g): mean	34.1	37.5	28.2
variance	25.7	26.8	10.0

* These values refer to the proportion of albino in coat.

small cells (S). In the other class of chimaeras, C is substituted for S. On rearrangement, this gives

$$G = S + (L - S)P = S + DP$$

where D is the difference in weight between the constituent strains. G is thus a linear function of P, as implied by Fig. 1. Since S, for present purposes, is a constant,

$$\text{var}(G) = D^2 \text{ var}(P)$$

where var(G) is the variance of G, and so on. Trivial reorganisation gives an expression for var(P)—the variance of proportions of cells in the growth-controlling tissue. D is directly observable and var(G) may be estimated, at least roughly, as follows. The variance in body weight of the C⟷C chimaeras does not

contain any var(G), by definition, since G was defined in terms of the proportions of differing cell types; however, the variance of the C⟷C chimaeras will reflect all other sources of variation. Subtracting the variance in body weight of C⟷C chimaeras from that of the S⟷L chimaeras therefore provides a direct estimate of var(G) appropriate to that class. All the necessary data are available from Table 1, and using the expression derived above, we obtain estimates of 0.046 and 0.190 for var(P) in S⟷L and C⟷L chimaeras, respectively. Viewed slightly differently, these are the variances in the cellular composition of the growth-controlling tissue that would be necessary to explain the increased variance in weight found in chimaeras between different strains. Although we make no claims for the numerical accuracy of these estimates of var(P), the values found seem to be reasonable, and they correspond closely to the variance in the proportions of melanocytes, estimated from the coat by direct observation (see Table 1).

Our results therefore establish that growth in chimaeras is a direct function of the proportions of cells from the two constituent strains, which in turn implies that growth regulation resides in the cellular properties of certain organs or tissues. Further, from consideration of the variances of the proportions of cells, it seems that the system of growth regulation needs be neither more complicated nor more extensive than that controlling the distribution of melanocytes in the skin. This indicates that the organs or tissues governing growth control may eventually be located; and to that aim we are marking constituent strains with variants of an enzyme that can be assayed quantitatively. The higher the correlation between body weight and proportion of L cells in various tissues, the more important will that tissue be in regulating growth. Equally, the systemic effect deriving from such tissue can be examined against the cellular composition of target organs, to see whether there is any local autonomy of size regulation.

R. C. Roberts
D. S. Falconer
Patricia Bowman
I. K. Gauld

ARC Unit of Animal Genetics,
Institute of Animal Genetics,
West Mains Road,
Edinburgh EH9 3JN, UK

Received October 17, 1975; accepted February 9, 1976.

[1] Bowman, P., and McLaren, A., *J. Embryol. exp. Morph.*, **23**, 693–704 (1970).
[2] Falconer, D. S., *Genet. Res.*, **22**, 291–321 (1973).
[3] Falconer, D. S., and Isaacson, J. H., *Genet. Res.*, **20**, 291–316 (1972).

Printed in Great Britain by Henry Ling Ltd., at the Dorset Press, Dorchester, Dorset

PAPER 16

Growth control in chimaeras.

Symp. on Genetic Mosaics and Chimaeras in Mammals. Gatlinburg, Tenn. pp. 39 - 49. Ed. L.B. Russell. Plenum Press. 1978.

by

D.S. FALCONER, I.K. GAULD & R.C. ROBERTS

From: GENETIC MOSAICS AND CHIMERAS IN MAMMALS
Edited by Liane B. Russell
(Plenum Publishing Corporation, 1978)

GROWTH CONTROL IN CHIMAERAS

D. S. Falconer, I. K. Gauld and R. C. Roberts

Agricultural Research Council, Unit of Animal Genetics,
West Mains Road, Edinburgh EH9 3JN, Scotland

How do chimaeras of large and small strains grow? Is their growth rate related to the cellular proportions in the body as a whole, or in any particular organ? Do the cells from the faster growing component tend to outgrow those from the slower growing? These are the main questions that we have been trying to answer, and this paper is a preliminary account of the results.

The strains of mice from which aggregation chimaeras were made were the Q-strains (Falconer, 1973), selected over 23 generations for large (L) and for small (S) body size, with unselected controls (C). There were six replicates each of L, C and S lines. Weight at 6 weeks of age was the criterion of body size. The Large lines were about twice the weight of the Small lines at this age. Luckily some of the replicates in each size group (L, C, and S) were polymorphic for the enzyme GPI-1 and for albino. In each size-group two stocks were constructed which were homozygous for complementary alleles at each of these two loci. Thus one stock was *Gpi-1*^{a} and albino to be referred to as (a), the other *Gpi-1*^{b} and coloured, to be referred to as (b). For this report we have data from 16 chimaeras of C(a) ↔ L(b) and 15 chimaeras of either S(a) ↔ L(b) or S(b) ↔ L(a). In addition to these overt chimaeras, there are a few single-colour animals in both the C ↔ L and S ↔ L groups, and a few chimaeras of types L ↔ L, C ↔ C and S ↔ S, to be used for some comparisons. The relation of body-size to the composition of the coat alone in the C ↔ L chimaeras and in a different set of S ↔ L chimaeras (not enzyme marked) was described by Roberts, Falconer, Bowman and Gauld (1976).

The chimaeras were scored for percent albino in the coat, and for the percent of allozyme *Gpi-1*^{a} in various tissues. The results

throughout are expressed as the percent of the Large component. The enzyme was scored by electrophoresis of serial dilutions, as described by Klebe (1975). The proportion of one or other allozyme can be derived from the number of dilution steps separating the last visible bands of the two. This method proved to be highly repeatable, nearly always giving the same reading on repeated runs. Nine organs or tissues were studied: they are listed in Table 1.

The C ↔ L chimaeras were killed for the enzyme assays when they were 10—12 months old, but their body weights analysed were those at 6 weeks of age. The enzyme content of the blood was

TABLE 1. Mean cell proportions (% L) in organs studied, ± standard errors

	C ↔ L	S ↔ L
Coat	49 ± 8	45 ± 5
Brain	46 ± 4	47 ± 4
Spinal cord	41 ± 6	46 ± 5
Pituitary	36 ± 8	50 ± 7
Liver	43 ± 7	58 ± 6
Lung	39 ± 7	53 ± 5
Kidney	45 ± 8	53 ± 5
Spleen	36 ± 6	55 ± 5
Blood	27 ± 7	53 ± 7
Mean*	40.6 ± 2.2†	50.9 ± 1.4
P‡	< 0.001	< 0.05

* Unweighted mean of organs.

† Significantly different from 50%, $P < 0.01$.

‡ Significance of variation between organs, from 2-way analysis of variance.

assayed at 6 weeks and at killing: there were no consistent changes. The S ↔ L chimaeras were all killed at 6 weeks of age, so that the weights and cell proportions refer to the same age. In both sets of chimaeras the 6-week weights of females were converted to male-equivalents by multiplying them by 1.2, a conversion factor found to apply equally to all three size-groups (Falconer, 1973). After conversion the sexes were pooled. All body weights were adjusted by regression to a standard litter size of 2 at birth.

We consider first the question of cell selection during development. Table 1 gives the mean cell proportions in each of the organs. In the C ↔ L chimaeras there were less than 50% of cells derived from the Large component in all organs, the overall mean of 41% Large being significantly different from 50%. The organs of the S ↔ L chimaeras varied round 50% with an overall mean of 51%. In both sets of chimaeras the organs were significantly heterogeneous with respect to cell proportions. These results answer one question clearly: there was no tendency for the cells from the larger component to outgrow those from the smaller; indeed the reverse was true in the C ↔ L chimaeras. There is, however, clear evidence of cell selection taking place to different degrees in different organs.

Fig. 1 gives a general impression of the relationship between weight and cell proportions. It plots body weight against the mean cell proportions in all the organs studied, which is the nearest we can get to the cell proportions in the body as a whole. It is very clear from both groups of chimaeras that body weight is influenced by the cell proportions. Both of the linear regressions shown on the graphs are significantly different from zero ($P < 0.001$). Estimates of the weights of the constituent strains are shown by arrows at the margins of the graphs. The C ↔ L chimaeras seem to show chimaeric heterosis to a marked degree. There are, however, difficulties in getting strictly comparable weights of the constituent strains, so the heterosis may be spurious. The S ↔ L chimaeras show no heterosis, and we think the evidence for chimaeric heterosis is not convincing.

The question now is: can we identify any organ as being more important than the others in influencing body weight? First consider the simple correlations between body weight and cell proportions in each organ. These are given in Table 2, arranged in descending order. The correlations are all high, ranging from 0.85 down to 0.65. The simple correlations, however, tell us very little, partly because the organs do not differ much, but mainly because the organs themselves are all highly correlated one with another in respect of cell proportions. This is illustrated in Table 3, which gives the distribution of correlations in each set of chimaeras. The high inter-organ correlations make it difficult

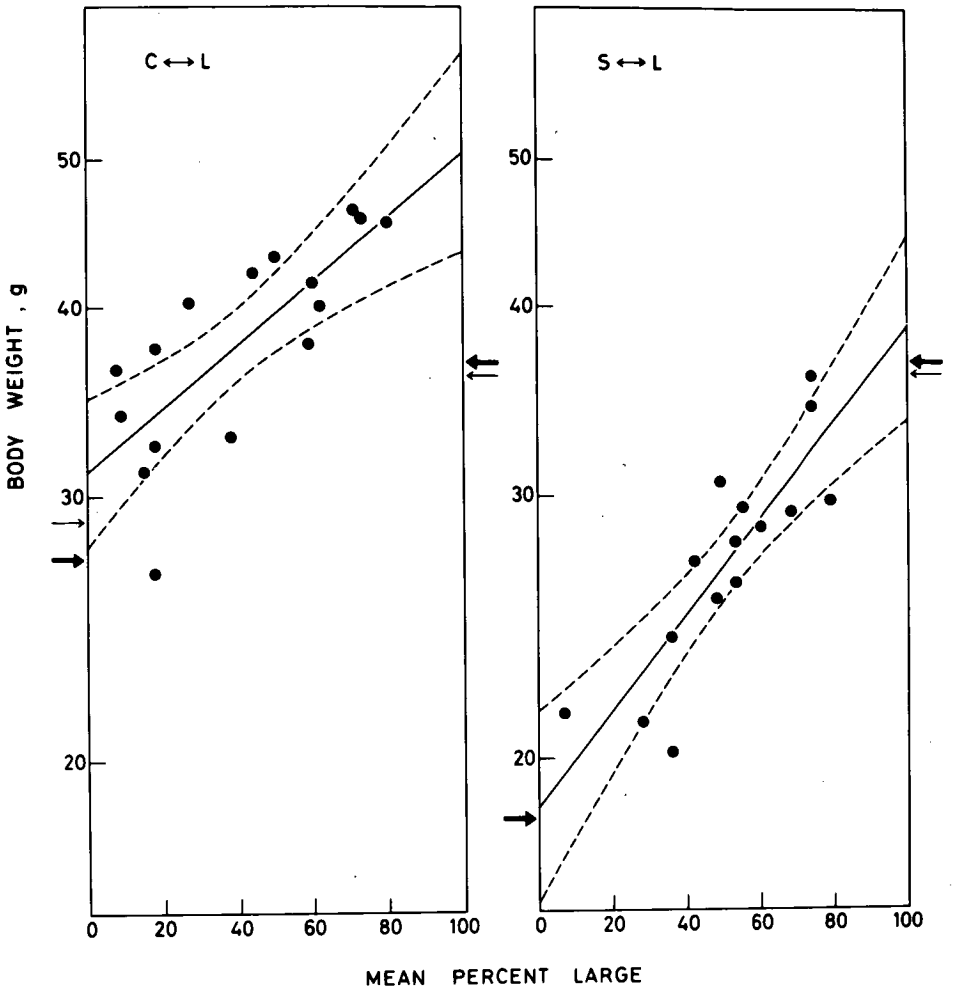


FIGURE 1. Relation of body weight at 6 weeks to cell proportions in nine organs or tissues. Weight is plotted on a logarithmic scale. The cell proportions are the means, unweighted, of all the organs. Each point is an individual chimaera. The straight lines are the fitted linear regressions. The broken lines are the 95% confidence limits of predicted mean weights. The arrows at the margins are estimated weights of the constituent strains as given by Roberts et al. (1976) (thick arrows), or from single colour chimaeras and L ↔ L or C ↔ C chimaeras (thin arrows).

TABLE 2. Simple correlations of body weight with % large cells in each organ, in order of magnitude of the correlation. *n* is the number of chimaeras.

C ↔ L (n = 16)		S ↔ L (n = 15)	
Coat	0.79	Pituitary	0.85
Pituitary	0.79	Spinal cord	0.84
Lung	0.77	Blood	0.79
Spinal cord	0.76	Kidney	0.78
Kidney	0.74	Lung	0.78
Blood	0.72	Brain	0.75
Liver	0.69	Spleen	0.72
Spleen	0.68	Coat	0.68
Brain	0.65	Liver	0.66
Mean	0.732 ± 0.051	Mean	0.761 ± 0.064

to separate their effects on weight. They also raise another question, which must be dealt with first. This is: are there real differences in cell proportions between the organs of the same mouse? Perhaps each mouse has its overall cell proportions from which its organs deviate only by errors of estimation. If there were no real differences between organs the question of whether any organ is more important than the others in determining weight could not be pursued further. The differences between organs within mice were, however, undoubtedly real. First, the differences were often very much greater than any found in repeat runs. Second, the error variance was calculated from the differences between left and right kidneys and between two samples of blood. Tested against this error variance, the mean square between organs within mice was highly significant ($P < 0.001$).

TABLE 3. Distributions of simple correlations between pairs of organs in respect of % large cells. The organs are the 9 listed in Table 1, giving 36 pairs.

correlation	number of organ-pairs C ↔ L	S ↔ L
.90 - .95	12	4
.85 - .90	6	7
.80 - .85	7	10
.75 - .80	10	7
.70 - .75	1	3
.65 - .70	0	2
.60 - .65	0	2
.55 - .60	0	0
.50 - .55	0	1
Mean	0.852	0.800
s.e.	± 0.066	± 0.092

To assess the effects on weight of each organ separately one would like, ideally, to calculate partial correlations. Unfortunately the number of animals is not much greater than the number of variables and so this approach is inpracticable. The alternative approach adopted is as follows. The organs (i.e. their cell proportions) are regarded as predictors of weight. When knowledge of all the organs is utilized, a certain proportion of the variance of weight is accounted for. The remainder, the residual variance, is attributable to environmental variance in the usual sense, together with any effects of other organs not studied. We first calculate the multiple correlation, R, of weight with all the organs. The residual variance of weight, as a proportion of the total, is $1 - R^2$. We then calculate the residual variance again with one organ omitted, and ask: is the residual variance now significantly

greater? In other words, does this omitted organ tell us anything more about weight beyond what all the other organs together tell us? This was repeated with each organ omitted in turn. Table 4 gives the results. The organs are arranged in order of importance

TABLE 4. Residual variance of weight, as percent of total, when one organ is omitted from the multiple correlation of weight with cell proportions in organs.

C ↔ L		S ↔ L	
Omitted	$1-R^2$ %	Omitted	$1-R^2$ %
None	16	None	8
Coat	30	Pituitary	25
Brain	29	Blood	20
Blood	25	Kidney	20
Spleen	18	Spinal cord	18
Lung	18	Liver	16
Liver	17	Spleen	14
Spinal cord	17	Brain	11
Pituitary	16	Lung	11
Kidney	16	Coat	8

	C ↔ L	S ↔ L
Minimum values for significance at P = 0.02 :	43	27
P = 0.05 :	32	19
P = 0.10 :	26	15

as judged by the increase of the residual variance that their omission causes. In the C ↔ L chimaeras no organ has a significant effect. In the S ↔ L chimaeras three organs, pituitary, blood and kidney, have effects significant at $P < 0.05$. The two series of chimaeras, however, are not at all consistent in the order of importance of the organs. For example pituitary, which has the biggest effect in the S ↔ L chimaeras, has no effect at all in the others. The results of this analysis therefore cannot be accepted as revealing any real differences between the organs in their effects on the control of growth. The proportion of the two cell types undoubtedly does affect weight, but we cannot localize the effect. There may of course be a localized effect in some other organ or tissue not studied. We ought therefore to ask whether the residual variance of weight contains any variance due to cell proportions that is not accounted for by the nine organs studied.

The variance of weight of chimaeras has three components: (1) that due to the differing cell proportions, which might be called 'chimaeric variance', arising from the genetic differences between the constituent strains, (2) that due to genetic differences between individuals within the constituent strains, and (3) that due to environmental differences affecting the chimaeras. There is no need here to separate (2) from (3) and they will be referred to jointly as environmental variance. The square of the multiple correlation, R^2, estimates the chimaeric variance that is accounted for by the nine organs, as a proportion of the total. The residual variance, $1 - R^2$, is the environmental variance together with any chimaeric variance that is not accounted for. The expected amount of environmental variance (components 2 + 3) can be estimated from the constituent strains or, better, from chimaeras made from strains of similar weight, and from single-component chimaeras. Comparison with the residual variance of the C ↔ L and S ↔ L chimaeras will then show whether there is any indication of chimaeric variance due to the cell proportions in any organ not studied. Table 5 gives the data for this comparison. The environmental variance is shown separately for the three genotypes (L, C, and S), each genotype being represented by different types of chimaera as shown at the foot of the table. The weighted mean of the environmental variance is 10.8 g^2, whereas the weighted mean of the residual variance is 9.5 g^2. The conclusion is that the cell proportions in the nine organs account for all the chimaeric variance. There is therefore no evidence of any other organ or tissue that controls growth. This does not mean that the existence of such an organ is excluded by the evidence. Suppose, for example, that connective tissue were the only controlling tissue, no other organ having any effect on growth. Being widely dispersed and probably with a large number of progenitor cells, connective tissue would be highly correlated in respect of cell proportions with most other organs. The meaning of the above result would thus be that the nine organs together give

TABLE 5. Environmental variance of weight (g^2) for comparison with residual variance.

Environmental*	d.f.	Variance
L genotype	9	21.2
C genotype	40	9.2
S genotype	3	0.7
Weighted mean	52	10.8
Residual		
C ↔ L	6	13.5
S ↔ L	5	4.6
Weighted mean	11	9.5

* Chimaera types in genotypes, with numbers of animals:

L genotype: L ↔ L (4 overt, 4 single colour);
S ↔ L (3 single).

C genotype: C ↔ C (24 overt, 14 single);
C ↔ L (5 single).

S genotype: S ↔ S (4 overt).

Variance calculated within groups in parentheses and then pooled.

Some of these chimaeras are from earlier series that were not enzyme-marked.

us a very accurate estimate of the cell proportions in connective tissue. Suppose, in contrast, that there was one growth-controlling tissue derived from very few progenitor cells. In that case, its

cell proportions would not be highly correlated with other organs, and the nine organs would give only a poor estimate of its cell proportions. The evidence rules out this possibility and we can conclude that growth is not controlled by any tissue derived from a very small number of progenitor cells.

Finally, are the sizes of any organs influenced by their own cell proportions? Organ weights are, of course, closely correlated with body weight. If an organ has a higher proportion of 'Large cells' than the rest of the body will it be disproportionately large? To answer this question we calculated the partial regression of organ weight on cell proportions in the organ, with body weight held constant, body weight being the weight at killing without adjustment for litter size. This was done for the brain, kidney, liver, lung, pituitary, spleen, and testis in both sets of chimaeras. There were three significant regressions (spleen, pituitary, testis) out of a total of fourteen, but the two sets of chimaeras were not consistent and we do not think that this is convincing evidence that cell proportions influence organ weights. Since the sizes of most organs are generally thought to be regulated by functional needs it seems unlikely that they would be influenced by their own cell proportions.

From all these results it looks as if growth and body weight are determined by the cellular genotype throughout the whole body, though not by localized effects on the growth of each organ individually.

SUMMARY AND CONCLUSIONS

Aggregation chimaeras were made from strains of mice differing in body size and marked by albino and an enzyme variant (GPI-1). The cell proportions — percent of cells from the larger of the two component strains — in each of nine organs or tissues were estimated by electrophoresis of serial dilutions or by visual scoring of the coat. The object was to look for relationships between body weight and cell proportions.

Body weight was very clearly correlated with the mean cell proportions in all the organs. The organs were all highly correlated with each other in respect of cell proportions, but there were real differences between organs within mice. There was no clear evidence that any one organ by itself had a significant effect on body weight. The nine organs jointly accounted for all the chimaeric variance, leaving no more than would be expected for environmental variance. There was no convincing evidence that the cell proportions in any organ had a localized effect on the weight of the organ itself.

The conclusions about the control of growth are: (1) The cells of the larger component do not proliferate faster than those of the smaller component during development. (2) None of the nine organs studied is predominant in controlling growth. (3) There may be some other organ or tissue that itself controls growth; but, if so, it cannot be one with a small number of progenitor cells. (4) Growth is correlated with the cellular genotype of each of the nine organs studied, either because each contributes something to the control of growth, or because each is correlated, in cellular genotype, with some other controlling organ. (5) If there is no other growth controlling organ, then it seems that growth depends on the cellular genotype throughout the whole body.

REFERENCES

Falconer, D. S. 1973. Replicated selection for body weight in mice. Genet. Res. 22: 291-321.

Klebe, R. J. 1975. A simple method for the quantitation of isozyme patterns. Biochem. Genetics 13: 805-812.

Roberts, R. C., D. S. Falconer, P. Bowman and I. K. Gauld. 1976. Growth regulation in chimaeras between large and small mice. Nature, Lond. 260: 244-245.

PAPER 17

The control of body size in mouse chimaeras

Genet. Res. 38, 25 - 46. 1981.

by

D.S. FALCONER, I.K. GAULD, R.C. ROBERTS & D.A. WILLIAMS

Genet. Res., Camb. (1981), **38**, *pp.* 25–46
With 2 text-figures
Printed in Great Britain

The control of body size in mouse chimaeras

By D. S. FALCONER, I. K. GAULD, R. C. ROBERTS
*Agricultural Research Council Unit of Animal Genetics**

AND D. A. WILLIAMS
Department of Statistics, University of Edinburgh

(*Received* 5 *November* 1980 *and in revised form* 17 *December* 1980)

SUMMARY

Aggregation chimaeras were made from embryos of strains of mice selected for large and small body size and of unselected controls. The strains were combined in pairs marked by albino coat colour and by allozyme variants at the *Gpi-1* locus. The proportion of cells derived from each component was scored visually in the coat melanocytes and by electrophoresis in ten other organs or tissues (blood, liver, lung, spleen, spinal cord, brain, pituitary, kidney, adrenal and testis). The object was to find out how body weight is related to cell proportions in the body as a whole and in the separate organs. Individuals varied widely in their mean cell proportions but there were significant differences between organs within individuals. Body weight was linearly related to the mean cell proportions which accounted for most, or possibly all, of the chimaeric variance of body weight. No one of the organs studied could be identified as being solely responsible for growth control, or as having a predominant influence on growth. The weights of some organs were probably influenced to a small extent by their own cell proportions independently of the individual's mean, but the differences of body weight were too great to be accounted for by the summation of localized effects on organs. The mean cell proportion, averaged over individuals, was close to 50%, proving that there was no tendency for cells from the larger component to outgrow those from the smaller. It is concluded that growth control must be systemic, but it was not possible to decide whether the systemic effect comes from some particular organ not studied, or is in some undefined way the consequence of the cell proportions in the body as a whole. There was some evidence, though it was inconclusive, that chimaeras show 'heterosis' for body weight.

1. INTRODUCTION

The aggregation chimaeras to be described in this paper were made by the fusion, or 'aggregation' of two 8-cell embryos derived from different strains. The resultant chimaeric mice contain two populations of cells, one derived from each of the constituent strains that provided the two embryos. Individual chimaeras

* Address: Institute of Animal Genetics, West Mains Road, Edinburgh EH9 3JN, Scotland.

differ widely in the relative proportions of the two cell populations in their bodies. This variation of cell proportions provides an opportunity to study the cellular control of any characteristics by which the two constituent strains differ. Nesbitt (1978) gives a preliminary account of a study of behavioural differences by this means. We made chimaeras of strains differing in body weight with the object of finding out how body weight is related to the cell proportions in the body as a whole and in some of the organs. If there is one particular organ that controls growth rate, we should expect to find body weight to be related to the cell proportions in this organ and not in any other organ. If there is a growth controlling organ or tissue that is derived from a single progenitor cell, we should expect to find a discontinuous distribution of body weight in the chimaeras, some being like one of the constituent strains and some like the other. If the cells of the constituent strains differ in their intrinsic rates of proliferation, we should expect the adult chimaeras to contain a higher proportion of cells from the faster growing strain. Finally, if the growth of individual organs is influenced by their own cells, we should expect to find organ weights to be related to the cellular proportions in that organ. These are the main questions that we set out to answer.

The strains from which the chimaeras were made were differentially marked by variants of the enzyme glucose phosphate isomerase, GPI, at the *Gpi-1* locus, so that the cell proportions in organs and tissues could be estimated by electrophoresis. They were also marked by albino so that the cell proportions in the coat melanocytes could be estimated visually. Some of the chimaeras have been described in two preliminary accounts (Roberts *et al.* 1976; Falconer, Gauld & Roberts, 1978*a*). All the chimaeras marked by the enzyme variant are included in the total of 87 chimaeras to be described here.

2. STOCKS USED AND CHIMAERAS OBTAINED

The mice used were from the Q-strains described by Falconer (1973). There were six replicate lines selected independently for large size, six selected for small size and six unselected controls. Selection was initially continued for 23 generations, after which it was relaxed. At generation 27 most of the lines were found to be polymorphic for the *Gpi-1* locus, and some were segregating for the albino gene, *c* (Garnett & Falconer, 1975). In order to construct stocks suitably marked for making the chimaeras, crosses were made between replicate lines at generations 34, 35, 42 and 43. Pairs of strains differing from each other at both marker loci were constructed from the large lines, from the control lines, and from the small lines, as shown in Table 1. In all three pairs of strains $Gpi\text{-}1^a$ was associated with albino (*c*) and $Gpi\text{-}1^b$ with coloured ($+^c$). After the strains had been made homozygous for their markers they were maintained by random mating. All the chimaeras to be described were made after homozygosis of the markers had been proved. They were made over a period of four years. The mean weights of the constituent strains over this period are given in Table 1.

Having two strains of each size meant that chimaeras from strains of different

sizes could be made reciprocally, and also that chimaeras could be made from strains of the same size. There were thus 9 possible types of chimaera. Of these, 8 are represented in the data, though one type by only 3 animals. Table 2 shows the types of chimaera, their designations and the numbers obtained.

Table 1. *Origins and body weights of the strains used, and markers made homozygous*

Q-lines crossed	Markers	Mean weight (g) males at 6 wks.
LD × LE	*Gpi-1*^{a} *c*	33·4
LB × LF	*Gpi-1*^{b} +	32·7
CA × CB	*Gpi-1*^{a} *c*	25·5
CD × CE	*Gpi-1*^{b} +	21·7
SA × SB	*Gpi-1*^{a} *c*	17·2
SF*	*Gpi-1*^{b} +	16·4

* This strain was not started from a cross, though subsequently a few animals from SE were introduced.

Table 2. *Numbers and types of chimaeras obtained*

				Chimaeras survived and used				
					Overt		Single-component	
Type	Embryos transferred	Chimaeras born	Success rate (%)	Total	♀	♂	Larger	Smaller
L/C	215	6	2·8	5	1	2	0	2
C/L	312	23	7·4	22	5	11	1	5
L/S	346	8	2·3	8	2	4	2	0
S/L	1110	31	2·8	26	4	15	5	2
S/C	122	4	3·3	3	1	2	0	0
L/L	181	8	4·4	6	3	0	3	
C/C	208	8	3·8	8	2	3	3	
S/S	132	9	6·8	9	2	6	1	
					20	43	8 7	9
Totals	2626	97	3·7	87	63		24	

The following terminology will be used. The chimaera types are designated by two letters referring to the sizes of the constituent strains, L for large, C for control and S for small. The first letter always refers to the strain marked by *Gpi-1*^{a} and albino. Thus L/C and C/L, for example, are reciprocal types of L↔C chimaeras, made from two different L-strains and two different C-strains. Chimaeras made from strains of the same size, as L/L, will be referred to as like-size chimaeras. Chimaeras displaying both cell populations in some part of the body will be referred to as overt chimaeras. Animals obtained from aggregated embryos but having only one cell population will be referred to as single-component chimaeras because, though not in fact chimaeric, they have been obtained by the same procedure and treatment. There were no animals that were non-chimaeric in the coat but chimaeric elsewhere in the organs studied.

Altogether there are 63 overt chimaeras for study, of which 47 were from strains differing in size and 16 were like-size chimaeras. The sex ratio among the overt chimaeras is not significantly different from the expected 75 % of males (McLaren, 1976). There are 24 single-component chimaeras, which will be used for various comparisons. The proportion of 28 % single-component chimaeras is in line with other studies (See Falconer and Avery, 1978, for a discussion of their origin).

3. METHODS

The method of aggregation followed was that described by Bowman and McLaren (1970). The host females to which the cultured embryos were transferred were mostly from the Control strains, though later females from the CFLP strain (Carworth, Europe) were used. At first, vasectomized males were used to induce pseudopregnancy in the host females; later the females were mated to entire males genetically marked by *Re Re*, and the chimaeras were then reared in litters with the progeny of the mating. The success rates in obtaining live young from aggregated embryos was rather low (Table 2). Dissection of host females that failed to produce litters proved that the losses of chimaeric embryos were almost all pre-implantation. The success rate of the C/L type is significantly higher than in the others. This may have been due to intrinsic properties of the C/L strain combination, but is more likely to have been due to unidentified technical factors because most of these chimaeras were made over a short period of time when no others were made.

The proportion of albino in the dorsal coat pigmentation was scored visually in 5 percent intervals. This was done at 3 weeks, 6 weeks, and when the chimaera was killed. It was also done again later, on the dried skins, for reasons that will be explained later. Most of the chimaeras were killed at 6 weeks of age, but those of one type, the C/L, were kept for breeding tests and were killed at 50 weeks. (The breeding tests conformed to expectation and are not described.) The organs studied are listed later, in Table 6. The carcasses were thoroughly drained of blood before removal of the organs. The proportions of the enzyme markers in the organs were estimated by electrophoresis of serial dilutions (Klebe, 1975), the electrophoresis being done by the method described by Shaw and Prasad (1970). The enzyme extracted from the smallest organs – pituitaries and adrenals – was barely enough for the serial dilutions. Consequently the assays of these organs were obtained from only some of the mice, and the cell proportions are less reliably estimated than those of the other organs. (Ovaries were also assayed but are not included in any of the analyses because too few records were obtained.) The cell proportions were estimated by the serial dilution method as follows.

Chimaeras contain a mixture of the two allozymes, and produce two bands on the gel, hybrid bands being absent from the tissues studied. The two allozymes have approximately the same specific activities (Padua, Bulfield & Peters, 1978), so the relative density of staining of the bands depends on the relative amounts of the two allozymes in the extract, and this in turn depends on the proportions of

the two cell populations in the tissue. Two dilutions of the extract are found which equalize the density of staining of the two bands, so that the amount of one allozyme in one dilution is equal to the amount of the other allozyme in the other dilution. The relative amounts of the allozymes in the original extract are then found from the dilution factor. Equality of staining is most easily judged by making a series of dilutions in equal steps and noting the dilutions at which each band just becomes invisible, i.e. the extinction points. Choice of the concentration of the initial extract and of the dilution factor depend on two things: (*a*) the number of dilutions that can be run in parallel on the same gel, and (*b*) the most extreme cell proportions that it is desired to quantify. The number of dilutions that could be run on the same gel was fifteen, and the most extreme cell proportions were taken to be 5% and 95%. By running artificial mixtures of the allozymes in the proportions 5:95, the initial concentration and the dilution factor were chosen so that the allozyme at 5% was visible in dilution-0 but not in

Table 3. *Scale of cell proportions*

(P%, estimated from the number of dilution steps, n, by which the extinction points differ, with a dilution factor of $r = 0{\cdot}8$).

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
P%	50·0	55·6	61·0	66·1	70·9	75·3	79·2	82·7	85·6	88·2	90·3	92·1	93·6	94·8	95·8

dilution-1, while the allozyme at 95% was visible in dilution-14 but not in dilution-15. By choice of an appropriate value of dilution-0 for each organ-extract, it was found that a dilution factor of 4/5 met these requirements, i.e. each dilution had a concentration of 0·8 of the next stronger one. The cell proportions are then obtained as follows. Let r be the dilution factor and n the number of dilution steps by which the extinction points differ. Then the proportion of cells giving the stronger band is $P = 1/(1+r^n)$. The cell proportions are not linearly related to the number of steps by which the extinction points differ. Consequently the method is more sensitive to differences of cell proportions at the extremes than it is near the middle of the range. Table 3 gives the cell proportions corresponding to the step differences. Tests of known mixtures of the allozymes showed that 50:50 mixtures had the same extinction points ($n = 0$), and that the method gave reliable readings for other mixtures. The small differences in activity noted by Padua *et al.* (1978) were therefore not enough to cause detectable error.

There were not enough chimaeras to justify analysing the sexes separately. Therefore for analyses involving body weights, except where otherwise stated, the weights of females were converted to male-equivalents by multiplying them by the factor 1·2, which was previously found to apply to the Q-strains generally (Falconer, 1973). In order to eliminate some of the environmental variation of body weight, adjustments were made for differences in the size of litter in which the chimaeras were born. The male-equivalent weights were adjusted by regression to a standard litter size of 4. The regression coefficient used was $b = -0{\cdot}59$ g per unit of litter size. This regression was estimated from the two control strains since most of the chimaeras were reared by females of these strains. Where mice of the

constituent strains are used for comparisons, their weights are adjusted to the same standard of 4, by the regressions estimated from the strains themselves, which were −0·87 for the two large strains, −0·59 for the controls, and −0·19 for the two small strains. The adjusted weights of the constituent strains contemporaneous with the chimaeras to which they gave rise are given in Table 4.

Table 4. *Mean 6-week weights (g) of the constituent strains for comparisons with the chimaeras*

(The weights are of males contemporaneous with the chimaeras and adjusted by regression to a standard litter size of 4.)

Chimaera type	Component L	Component C	Component S	Difference
L/C	35·35	24·38	—	10·97
C/L	41·32	28·88	—	12·44
L/S	35·96	—	16·32	19·64
S/L	37·98	—	17·30	20·68
S/C	—	23·01	17·79	5·22
	1st	2nd	Mean	
L/L	37·86	41·22	39·54	
C/C	28·70	23·25	25·98	
S/S	17·97	17·97	17·97	

4. RESULTS

(i) *Cell proportions*

There are several questions to be answered about cell proportions in overt chimaeras before body weight is brought into consideration. The main question is whether there is any tendency for cells from the larger constituent strain to outgrow those from the smaller. Other questions concern the distributions of cell proportions, the correlations between organs, and whether there are real differences in proportions between organs of the same individual.

Mean cell proportions. Let P be the proportion of cells from the larger strain ('large' cells for short), in any organ of an individual. The mean of all the organs measured in a particular individual estimates the cell proportion in the body as a whole. This will be referred to as the mean cell proportion and symbolized by $\bar{P}$. The mean cell proportion of an individual, $\bar{P}$, will first be taken as the unweighted mean P of all the organs measured, paired organs being averaged. Table 5 shows the mean value of $\bar{P}$ averaged over the individuals in each chimaera type. None of the means is significantly different from 50 %, nor are the chimaera types significantly different from each other. The conclusion is therefore clear that, when averaged over all organs, there is no tendency for one cell population to outgrow the other.

Table 6 shows the cell proportions, P, in each organ separately, averaged over all individuals. It also shows the variance of P about 50 % and the pooled variance within chimaera types. A significant reduction of the within-type variance, when

Table 5. *Mean cell proportions in overt chimaeras*

(The values tabulated are the means of $\bar{P}$ %, where $\bar{P}$ is the mean cell proportion of all organs of an individual. The standard errors are derived from the variance of $\bar{P}$ among individuals. For the L/L, C/C and S/S types, $\bar{P}$ is the proportion of cells from the albino, *Gpi-1*^{a} component; for all the other types it is the proportion of cells from the larger component strain. The standard errors are based on the pooled variance within chimaera-types.)

Chimaera type	number of individuals	Mean of $\bar{P}$, %, ± S.E.
L/C	3	28·6 ± 12·1
C/L	16	41·2 ± 5·2
L/S	6	53·2 ± 8·6
S/L	19	49·0 ± 4·8
S/C	3	45·5 ± 12·1
L/L	3	37·1 ± 12·1
C/C	5	52·6 ± 9·4
S/S	8	61·0 ± 7·4

(Note: The standard errors of the C/L and S/L types given in Falconer, Gauld & Roberts (1978*a*) were inappropriate, and the conclusion that C/L differed significantly from 50 % was wrong.)

Table 6. *Means, variances about* 50 %, *and within type variances of the proportions of* Gpi-1^a *in all overt chimaeras*

	All overt chimaeras ignoring type				
		Variance about 50 %		Within type	
Organ	Mean ± S.E.	D.F.	Variance	D.F.	Variance
Coat	45·9 ± 3·4	63	715	55	655
Blood	46·4 ± 3·9	63	939	55	857
Liver	50·9 ± 3·1	62	581	54	572
Lung	49·7 ± 3·0	63	559	55	561
Spleen	47·1 ± 3·0	61	559	53	491
Sp. Cord	44·9 ± 3·0	47	448	39	440
Brain	47·9 ± 2·1	63	272	55	265
Pituitary	46·4 ± 3·5	50	627	42	444
Kidney	48·1 ± 2·8	63	503	55	487
Adrenal	49·8 ± 3·5	46	552	39	509
Testis	51·3 ± 4·0	43	684	36	712
Mean $\bar{P}$	47·5 ± 2·7	63	452	55	441

compared with the variance about 50 %, would suggest that there were differences between type means, or between the overall mean and 50 %. In fact the reduction is significant only in the pituitary ($P < 0{\cdot}01$). Here the differences between pituitary means are largely attributable to the configuration of missing pituitary values which causes the means of two types, L/L and S/S, to take extreme values. We regard this as a chance effect and proceed under the assumption that for each organ and each chimaera type the cell proportion is distributed about a mean of 50 %.

Distribution of cell proportions. Several series of chimaeras have shown that the

cell proportions in the coats have a more or less uniform distribution, all proportions in overt chimaeras being about equally frequent (Falconer & Avery, 1978). In this respect chimaeras differ from X-inactivation mosaics, which are much less variable in cell proportions in the coat. Falconer & Avery (1978) showed how a uniform distribution could arise from the sampling of cells to form the primary ectoderm in embryos of chimaeras. We have examined the distributions of cell proportions in the organs studied here, to see if they also show uniform distributions. In χ^2 goodness of fit tests, using 8 class intervals of width 12·5 %, there were significant ($P < 0{\cdot}05$) deviations from uniformity in six organs, namely liver, lung, spleen, spinal cord, brain and kidney. Moreover Table 6 shows that only the blood has a variance which exceeds the theoretical value of 833 for a uniformly distributed percentage. We conclude that the cell proportions in most organs have a distribution which is more concentrated around 50 % than the uniform distribution. The theory of Falconer & Avery (1978) therefore cannot be right in all details. To account for the distributions found here, however, it is only necessary to suppose that some cell mixing occurs before the separation of the primary ectoderm from the primary endoderm. Note that the variances in Table 6 do differ appreciably between organs, and that the brain, surprisingly, has smaller variance than the mean $\bar{P}$. We shall return to this point later in this section.

Correlations of cell proportions. The cell proportions in the organs of individuals are highly correlated. Table 7 gives the simple correlations of all organ-pairs in all the overt chimaeras irrespective of type. The correlations range from 0·37 to 0·89, and the average is 0·73. There are no obvious differences among the organs in the mean level of their correlations with other organs, except the testis which is on average clearly less highly correlated than other organs. The fact that blood does not show a higher than average correlation with other organs shows that contamination by blood has probably not introduced any serious error.

The left- and right-hand members of the paired organs are much more highly correlated than are different organs, all the correlations being over 0·9. These correlations are shown in the diagonal of Table 7. The coat was treated as a 'paired organ' in the following way. As mentioned under Methods, the coats were rescored from the dried skins. As far as possible the scoring was based on the dorsal part of the coat in order to correspond with the scoring of the live animal. The skins of all the overt chimaeras were examined in turn four times, scoring first the whole dorsal coat, second the left half, third the right half, and finally the whole coat again. The repeated whole-coat scores will be used in the next section. The left-side and right-side scores are used as a paired organ.

A question of interest is whether organs on the same side of the body are more highly correlated than organs on different sides. This possibility was tested by comparing ipsi-lateral and contra-lateral correlations among the paired organs. Ipsi-lateral correlations were calculated from, for example, left kidney with left adrenal and right kidney with right adrenal, these two correlations being then averaged. Contra-lateral correlations were the average of left kidney with right adrenal and right kidney with left adrenal. There were 6 such inter-organ correla-

tions among the 4 paired organs. The ipsi-lateral correlations were very slightly, but non-significantly, greater than the contra-lateral; the mean difference being $0{\cdot}005 \pm 0{\cdot}005$. There seems, therefore, to be little or no tendency for the cell proportions to differ on the two sides of the body as a whole.

Components of variation in organ cell proportions. The large positive correlations between the cell proportions of all pairs of organs indicate that the main component of variation between individuals is a variable which represents the individuals' mean cell proportions. The unweighted mean, $\overline{P}$, could be used to represent this

Table 7. *Simple correlations between pairs of organs in respect of cell proportions*

(On the diagonal, correlations between left and right sides of paired organs. The mean at the foot of the table is the unweighted mean of the 10 correlations of each organ, excluding left-right correlations. (n) is the number of individuals with records of the organ.)

Organ	(n)		1	2	3	4	5	6	7	8	9	10	11
Coat	(63)	1	0·92	—	—	—	—	—	—	—	—	—	—
Blood	(63)	2	0·70	—	—	—	—	—	—	—	—	—	—
Liver	(62)	3	0·63	0·66	—	—	—	—	—	—	—	—	—
Lung	(63)	4	0·70	0·85	0·81	—	—	—	—	—	—	—	—
Spleen	(61)	5	0·70	0·86	0·72	0·83	—	—	—	—	—	—	—
Sp. Cord	(47)	6	0·77	0·81	0·74	0·88	0·77	—	—	—	—	—	—
Brain	(63)	7	0·74	0·74	0·73	0·80	0·76	0·85	—	—	—	—	—
Pituit.	(50)	8	0·68	0·80	0·68	0·79	0·82	0·82	0·80	—	—	—	—
Kidney	(63)	9	0·77	0·80	0·82	0·89	0·80	0·87	0·84	0·79	0·92	—	—
Adrenal	(46)	10	0·71	0·76	0·70	0·74	0·78	0·70	0·80	0·85	0·81	0·93	—
Testis	(43)	11	0·54	0·45	0·59	0·52	0·47	0·37	0·58	0·62	0·63	0·72	0·97
Mean			0·69	0·74	0·71	0·77	0·75	0·76	0·76	0·76	0·80	0·76	0·55

variable. However, the variances of $P-\overline{P}$, given in the first column of Table 8, differ substantially between organs and suggest that an average is better estimated iteratively by the weighted mean, $\overline{P}_w$, with weights inversely proportional to $\mathrm{var}(P-\overline{P}_w)$. These variances, $\mathrm{var}(P-\overline{P}_w)$, are given in the second column of Table 8. The values of the weighted mean, $\overline{P}_w$, are very similar to those of the unweighted mean, $\overline{P}$. The mean of $\overline{P}_w$ is $47{\cdot}7 \pm 2{\cdot}6$, the correlation between $\overline{P}_w$ and $\overline{P}$ is 0·994 and the variance of $\overline{P}_w - \overline{P}$ is 5·75.

It would be natural to represent an organ cell proportion P as the sum of two components, the individual's mean, $\overline{P}_w$, and the organ deviation $P-\overline{P}_w$. This representation would be particularly useful if the two components, $\overline{P}_w$ and $P-\overline{P}_w$, were independent. That such independence cannot be assumed is shown by the correlations between $P-\overline{P}_w$ and $\overline{P}_w$ given in the third column of Table 8. Two of the correlations are clearly significant, blood with $+0{\cdot}41$ and brain with $-0{\cdot}63$. The meaning of these correlations can be stated as follows. If an individual has its mean cell proportion above 50 % then the blood tends to be above the mean and the brain below; conversely an individual with mean below 50 % tends to have its blood below the mean and its brain above. Or, in other words, over all individuals blood tends to deviate more from 50 % than the mean does, while brain deviates less. In consequence the blood has a higher variance than the mean and the

brain has a lower variance, as was noted earlier in this section, and shown in Table 6. The relationship between deviation and mean must obviously be non-linear because the deviation $(P-\bar{P}_w)$ must be zero at three points, when $\bar{P}_w$ is 0, 50 and 100%.

Table 8. *Estimates of unexplained variation and the parameters of models relating cell proportions of individual organs to their average*

				Logistic regression of P on $\bar{P}_w$	
Organ	Var $(P-\bar{P})$	Var $(P-\bar{P}_w)$	Correlation of $P-\bar{P}_w$ with $\bar{P}_w$	$\beta \pm$ S.E.	Residual variance
Coat	221	245	0·06	1·05 ± 0·12	246
Blood	225	247	0·41	1·55 ± 0·14	203
Liver	163	160	0·03	1·00 ± 0·10	162
Lung	83	65	0·25	1·13 ± 0·06	63
Spleen	119	121	0·00	1·00 ± 0·09	123
Sp. Cord	91	67	−0·27	0·90 ± 0·06	64
Brain	96	82	−0·63	0·66 ± 0·04	51
Pituitary	122	143	0·20	1·18 ± 0·11	141
Kidney	75	55	0·09	1·05 ± 0·06	55
Adrenal	109	140	0·16	1·13 ± 0·12	139
Testis	342	404	−0·06	0·92 ± 0·19	413

A simple empirical model, which satisfies these constraints and gives rise to the observed correlations, is one that relates $E(P)$ to $\bar{P}_w$ linearly on the logistic scale. If $E(P)$ is the conditional expectation of P given $\bar{P}_w$ the model is:

$$\log \frac{E(P)}{100-E(P)} = \beta \log \frac{\bar{P}_w}{100-\bar{P}_w}.$$

Here the coefficients β vary between organs. A value of β less than unity corresponds to a negative correlation between $P-\bar{P}_w$ and $\bar{P}_w$, a value of β greater than unity corresponds to a positive correlation. Estimates of the β, and the residual variances of the $P-E(P)$, are given in the final two columns of Table 8. We shall not consider this model further except to note that an equally good empirical fit is obtained by the simpler linear regression

$$E(P)-50 = \beta(\bar{P}_w-50)$$

This has implications for the later study of the dependence of body weight on the components $\bar{P}_w$ and $P-\bar{P}_w$. Models of development that could give rise to the observed correlations of $P-\bar{P}_w$ with $\bar{P}_w$ will be considered in the Discussion.

Comparison of residual organ cell proportion variation with assay error. There would be no point in studying the deviations $P-\bar{P}_w$ further unless they represented real organ differences within mice rather than just assay error. To confirm that the $P-\bar{P}_w$ represent real organ effects we need to compare their variances with the error-variance due to errors of estimation of cell proportions. The enzyme-assays were not replicated, but there are nevertheless three ways by which their

error variance can be estimated. These estimates are all biased upwards, so that the significance of differences between organs will be under-estimated. The sources of the estimates are the following. (1) The spinal cord was divided into three roughly equal parts and each part was assayed separately but on the same day. The bias comes from differences in cell proportions between the three parts. (2) The blood was assayed on two occasions – at 6 and 50 weeks in the C/L chimaeras and at 3 and 6 weeks in all the others. The bias comes from changes in cell proportions with time. (3) Paired organs – kidney, testis, adrenal – were assayed separately. The bias comes from real differences between left and right organs.

The cell proportions in the coat need not have the same error variance as the other organs because they were measured differently. For the error variance of the coat scores we have three estimates:

(1) Measurements on the live animals repeated at 3 weeks and 6 weeks,

(2) Two repeated measurements on skins, as described earlier, and

(3) The left and right sides of skins.

Table 9. *Estimates of the error variance of percentage cell proportions*, P

	D.F.	Variance
Enzyme assays		
1. Spinal cord	94	4·77
2. Blood	61	49·23
3. Kidneys	60	36·08
Testes	42	18·15
Adrenals	38	37·84
Coat score		
1. 3–6 weeks	63	0·79
2. Whole skin	62	8·68
3. Half skins	62	45·97

Table 9 gives the various estimates of error variance in enzyme proportions and in coat scores. The estimates are all substantially smaller than the corresponding organ variances in Table 8, except in the case of the kidney where the difference is smaller but is significant ($P < 0{\cdot}05$). For these organs at least we may conclude that the $P - \bar{P}_w$ do measure real organ differences.

(ii) *Weight and cell proportions*

In this section we shall deal first with the relation between body weight and mean cell proportions, and then with its relationship with the cell proportions in separate organs.

Mean cell proportion. Fig. 1 shows the relationship between 6-week body weight and weighted mean cell proportion, $\bar{P}_w$, in the two chimaera types with the largest numbers. Body weights are male-equivalents adjusted for litter size. The mean weights of the contemporary constituent strains, from Table 4, are shown by arrows in the margins. It is very obvious that the weights are strongly influenced by the mean cell proportions. The calculated linear regressions of weight on $\bar{P}_w$

are shown with the 95 % confidence limits of predicted mean weights. The various parameters estimated from the regression analyses are given in Table 10.

In order to combine chimaeras of all types into a single analysis, weights were scaled to a standard difference between the constituent strains. For each chimaera a 'relative weight', w, was calculated as follows,

$$w = \frac{W - S}{L - S}$$

where W is the actual weight, and L and S are the mean weights of the larger and smaller constituent strains respectively. Where cell proportions are expressed as

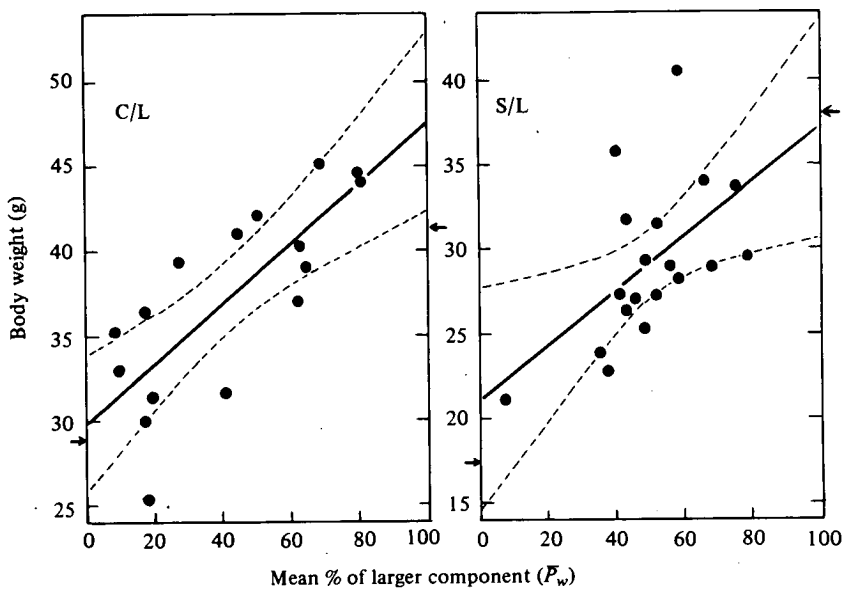


Fig. 1. Body weight at 6 weeks in relation to weighted mean proportion, $\bar{P}_w$, of cells from the larger component. Each point represents one individual. The arrows at the margins show the mean weights of the contemporaneous constituent strains. The lines are the fitted linear regressions of weight on $\bar{P}_w$, with 95% confidence limits of predicted mean weight.

Table 10. *Regression analysis of body weight* (W) *on mean cell proportion,* $\bar{P}_w$ *in percentage units, for C/L and S/L chimaera types*

	C/L	S/L
Number of mice	16	19
Correlation, W with $\bar{P}_w$	0·77	0·55
Strain difference, D	12·44 g	20·68 g
Regression, W on $\bar{P}_w$	0·177 ± 0·039	0·161 ± 0·058
Intercept at $\bar{P}_w = 0$	29·8 ± 1·9	21·0 ± 3·1
Intercept at $\bar{P}_w = 100$	47·6 ± 2·5	37·1 ± 3·1
Variance of $\bar{P}_w$	639·9	260·3
Total variance of W	33·89	22·05
Residual variance of W	14·65	16·29
σ_E^2 (Table 12)	10·36	10·36
Chimaeric variance not accounted for by $\bar{P}_w$	4·29	5·93

percentages, relative weights are also expressed as percentages, which means that the $L-S$ difference is standardized to 100. We assumed that the unexplained variation in weight W about any regression on organ cell proportions has a constant variance σ_w^2 over all chimaera types. Thus the variance of the relative weight w is $\sigma_w^2 (L-S)^{-2}$ which varies between chimaera types. All regressions involving w as the dependent variable were therefore estimated by weighted least squares using weights $(L-S)^2$.

The relationship in all chimaeras between relative weight w and mean cell proportion $\bar{P}_w$ is shown in Fig. 2. Estimates of linear regressions within chimaera

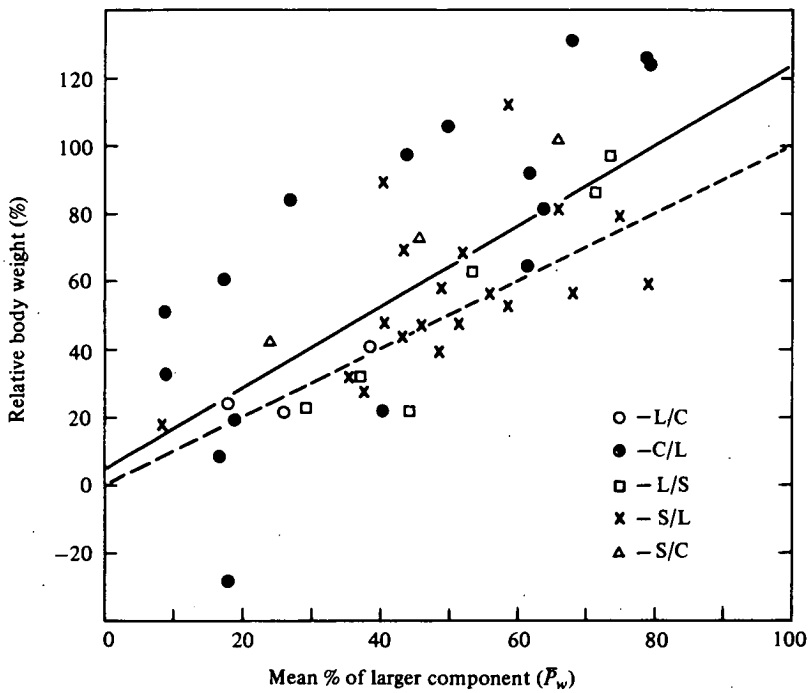


Fig. 2. Relative body weight in relation to weighted mean cell proportion, $\bar{P}_w$, both in percentage units. The continuous line is the fitted linear regression; the broken line joins the means of the constituent strains. Each point represents one individual.

Table 11. *Linear regressions of relative body weight* w *on mean cell proportion* $\bar{P}_w$, *in percentage units, for all five chimaera types*

Chimaera type	Intercept at $\bar{P}_w$ = 50% ± S.E.	Slope	Residual variance of Weight σ_w^2
L/C	48·8 ± 55·5	0·89 ± 2·32	
C/L	79·1 ± 7·8	1·43 ± 0·30	
L/S	50·9 ± 7·7	1·75 ± 0·47	13·55
S/L	56·9 ± 4·1	0·78 ± 0·26	(pooled within types)
S/C	79·6 ± 42·4	1·41 ± 2·37	
Combined	59·6 ± 3·3	1·08 ± 0·18	14·59

types, and of the overall regression ignoring chimaera type, are given in Table 11. Inspection of both Fig. 2, and intercept estimates in Table 11, suggest that the regression in the C/L chimaera type has a higher elevation than in other chimaera types. However, the F statistics for differences between slopes, and for differences between intercepts assuming a common slope were both non-significant. If there are any differences in elevation they probably resulted from errors in the strain means from which the relative weights were calculated.

The conclusions that can be drawn from the regression coefficients are rather limited. The regression lines must pass through the weight of the larger strain at $\bar{P}_w = 100\,\%$ and of the smaller strain at $\bar{P}_w = 0\,\%$. The expected linear regression of weight on cell proportion in percentage units is therefore D/100, where D is the difference in weight between the constituent strains; and the expected linear regression of relative weight on $\bar{P}_w$ is 1. The slope of the fitted overall regression agrees well with this expectation, but the intercept at $\bar{P}_w = 50\,\%$ is significantly higher than the expected value of 50. There is no evidence from a graphical study of the residuals that the regression is non-linear. We can therefore conclude that the body weight is linearly dependent on the cell proportion in some organ or organs.

The variance of weight and its partitioning by the regression analysis is more informative than the regression coefficients themselves. The regression on $\bar{P}_w$ partitions the variance as

$$\text{total variance} = (\text{variance due to } \bar{P}_w) + (\text{residual variance}).$$

An alternative partition is

$$\text{total variance} = (\text{'chimaeric variance'}) + (\text{'non-chimaeric' variance})$$

where 'chimaeric variance' is associated with differences in cell proportions and 'non-chimaeric' variance is the remainder not associated with cell proportions. The non-chimaeric variance is mainly environmental, but it contains also a component due to genetic differences within the constituent strains because individual chimaeras made from the same two strains will not have exactly the same genotypic values of their component embryos. This is non-chimaeric variance because it is not associated with differences of cell proportions. The two components of the non-chimaeric variance, however, do not need to be distinguished in what follows. The two partitions of total variance are not identical because the variance due to $\bar{P}_w$ forms part, but not necessarily all, of the chimaeric variance. There may be components of chimaeric variance not attributable to $\bar{P}_w$. In particular there may be components of chimaeric variance attributable to regressions on the deviations $P - \bar{P}_w$ of organs which influence body weight. We wish to know whether such components can exist, for it is only by identifying them that we shall be able to identify organs which influence body weight. By comparing

$$\begin{aligned}\text{total variance} = {} & (\text{chimaeric variance due to } \bar{P}_w) \\ & + (\text{chimaeric variance not due to } \bar{P}_w) \\ & + (\text{non-chimaeric variance})\end{aligned}$$

with the regression partition we deduce that

$$\text{chimaeric variance not due to } \bar{P}_w = (\text{residual variance}) - (\text{non-chimaeric variance}).$$

The possible existence of chimaeric variance attributable not to regression on $\bar{P}_w$, but to regression on the $P - \bar{P}_w$, can therefore be inferred from comparison of estimates of residual and non-chimaeric variances.

Table 12. *Non-chimaeric variance of body weight, estimated from like-size and single-component chimaeras*

Genotype	D.F.	Variance, g^2
Large	12	12·99
Control	13	8·82
Small	9	9·09
Pooled	34	10·36

Table 13. *Regressions of relative body weight,* w *on proportion of 'large' cells in each organ. For each organ* w *is regressed in turn on* P *alone, on* $\bar{P}_w$ *alone, on both* $\bar{P}_w$ *and* P.

Organ	Number of individuals	Linear regression w on $P \pm$ S.E.	Residual variances of W about regression on: P alone	$\bar{P}_w$ alone	both $\bar{P}_w$ and P
Coat	47	0·78 ± 0·15	16·10	14·59	14·30
Blood	47	0·68 ± 0·13	16·11	14·59	14·80
Liver	46	0·63 ± 0·17	18·74	13·94	13·98
Lung	47	0·85 ± 0·16	16·40	14·59	14·87
Spleen	45	0·80 ± 0·18	17·85	14·10	14·34
Sp. Cord	39	0·99 ± 0·18	13·93	13·07	13·36
Brain	47	1·22 ± 0·26	17·79	14·59	14·87
Pituitary	38	0·77 ± 0·18	17·47	13·94	14·21
Kidney	47	0·95 ± 0·16	14·77	14·59	14·73
Adrenal	33	0·64 ± 0·17	14·51	11·29	11·64
Testis	34	0·36 ± 0·18	21·56	14·77	14·80

An independent estimate of the non-chimaeric variance comes from single-component chimaeras and from overt chimaeras of like-sized strains. These chimaeras contain cells of a single size-genotype, that is to say the origins of their cells are either all Large, or all Control or all Small. Their variance represents all of the non-chimaeric variance, including any variance due to genetic differences within the constituent strains. Table 12 gives the estimates of the non-chimaeric variance derived from each of the three size-genotypes. The estimates do not differ significantly by Bartlett's test, so they are pooled to give a joint estimate of 10·36 g^2. The residual variance about the regression on $\bar{P}_w$ was estimated as 14·65 and 16·29 respectively in the C/L and S/L chimaera types (Table 10), and as 14·59 over all chimaera types (Table 11). None of these estimates differ significantly from the estimated non-chimaeric variance. Thus although there may be components of chimaeric variance not accounted for by $\bar{P}_w$, whose sum is best

estimated as $14{\cdot}59 - 10{\cdot}36 = 4{\cdot}2$, there is no significant evidence that such components exist.

Separate Organ Cell Proportions. It has been established that each organ cell proportion, P, may be expressed as the sum of two components $\bar{P}_w$, which is common to all organs, and the organ deviation $P - \bar{P}_w$ which depends on the organ but is approximately linearly dependent on $\bar{P}_w$ in some organs. The fact that the body weight is linearly dependent on $\bar{P}_w$ establishes a dependence on the cell proportions in one or more organs. To associate this dependence with a particular organ the dependence of body weight on the deviation $P - \bar{P}_w$, additional to the dependence on $\bar{P}_w$, must be established. An additional dependence of body weight on $P - \bar{P}_w$ would be indicated either by P being a better predictor of body weight than $\bar{P}_w$, or by a partial regression of body weight on P, eliminating $\bar{P}_w$.

Table 13 gives the statistics derived from simple regressions of relative body weight on P alone and on $\bar{P}_w$ alone, and from the multiple regression on $\bar{P}_w$ and P together. The main interest lies in the residual variances about these regressions: the lower the variance, the better is the predictor of body weight. The results are easily summarized: (1) for each organ, the organ itself (P) is a poorer predictor than the mean ($\bar{P}_w$). (2) For each organ, the addition of $\bar{P}_w$ to P gives a better prediction than P alone, and the decrease of residual variance is significant ($P < 0{\cdot}05$) for all organs except spinal cord and kidney. (3) For each organ, the addition of P to $\bar{P}_w$ does not give a significantly better prediction than $\bar{P}_w$ alone. In fact the residual variances are increased for all organs except the coat for which, however, the decrease is not significant.

Two conclusions can be drawn from these results. First, from (2), no one of these organs, with the possible exception of the spinal cord and kidney, can be solely responsible for determining body weight. Second, from (3), no one of these organs can be identified as playing a predominant role in determining body weight. The question of whether there may be some other organ that is the principal growth-controlling organ will be considered in the Discussion.

(iii) *Organ weights*

Is the weight of an organ influenced in any degree by the cells that it contains, independently of the cells in the rest of the body? For example, if a liver contains a higher proportion of 'large' cells than the rest of the body, will it be relatively larger in consequence? To answer this question we calculated the partial regression of organ weight on cell proportions in the organ with mean cell proportions held constant. The organ weights were first adjusted to a standard body weight. For this purpose weights were transformed to logs because log organ weights had previously been found to be linearly related to log body weights (Falconer, Gauld & Roberts, 1978*b*). For each organ, regressions of log organ weight on log body weight were calculated from all the chimaeras of all types. These regressions were then used to adjust the organ weights of overt chimaeras to a standard body weight. The kidneys of females were adjusted to male equivalents since the relative weights of female and male kidneys had been found to differ (Falconer *et al.*

1978*b*). The regressions of these adjusted log organ weights were then calculated from all overt chimaeras of unlike-sized strains combined, the regression being the partial regression on the proportion of cells from the larger component in the organ, with the mean cell proportion, $\bar{P}$, held constant. (The unweighted mean was used for these calculations.) The results for eight organs are given in Table 14. None of the regressions is significantly different from zero, though the liver and kidney are close to significance ($P \sim 0{\cdot}06$). Seven of the eight regressions are positive, which by a sign-test has a probability of $P = 0{\cdot}07$. The organs do not differ significantly in their regressions. The regression pooled within organs is not significantly different from zero ($P \sim 0{\cdot}2$). The pituitary and adrenal might be excluded on the grounds that their small size makes their records unreliable. Omitting them the pooled regression is still not quite significant ($P = 0{\cdot}065$).

Table 14. *Partial regressions of* log_{10} *organ-weight (adjusted to standard body weight) on proportion of 'large' cells,* P, *in the organ, with mean cell proportion,* $\bar{P}$ *held constant*

Organ	D.F.	Regression ± S.E.	*t*	% effect†	Cell mass diff. %‡
Spleen	42	0·238 ± 0·241	0·99	73	24
Kidney	43	0·198 ± 0·103	1·93	58	44
Brain	44	0·132 ± 0·091	1·45	36	—
Liver	43	0·111 ± 0·057	1·95	29	57
Testis	31	0·067 ± 0·178	0·38	17	—
Lung	44	0·060 ± 0·093	0·64	15	17
Adrenal	29	0·057 ± 0·330	0·17	14	—
Pituitary	35	−0·234 ± 0·301	0·78	−42	—
Pooled (1)*	325	0·076 ± 0·061	1·24	19	
Pooled (2)*	257	0·105 ± 0·057	1·84	27	

* Pooled within organs, (1) all organs, (2) with adrenal and pituitary excluded.

† 100 (antilog of regression − 1). See text for explanation.

‡ Percentage difference, L − S, in cell size, from Falconer, Gauld & Roberts (1978*b*).

Though not conclusive, the evidence does point fairly strongly to a local influence of cell proportions on the size of some organs. A localized effect on organ size is made more credible by the fact that the shapes of vertebrae are influenced by their cellular composition (Moore & Mintz, 1972). If the effect on organ weight is real, could its summation over all organs account for the effect on body weight as a whole? The ratio of body weights of the constituent strains was, on average, about L/S = 1·82, so that the larger was about 82 % heavier than the smaller. The comparable percentage differences produced by the localized effects on organs are shown in the column headed '% effect' in Table 14. The meaning of these '% effects' on organs is this. If two individuals have the same $\bar{P}$ and one had, say, a liver with all 'small' cells while the other had its liver with all 'large' cells, then the latter liver would be 29 % heavier than the former. So, if all organs had the same localized effect as the liver, and there were no effect of a growth-controlling organ, a mouse with all 'large' cells in its body would be 29 % heavier

than one with all 'small' cells. This is much less than the actual difference of 82 %. The localized effects on the liver and the lung are significantly less than 82 %, and so are the two pooled effects. It is therefore very unlikely that the overall body weight is determined simply by the summation of localized effects on organ weights.

The localized effects of cellular genotype, if real, could be the consequence of differences of cell size. The cell size (mass of organ per nucleus) in four of these organs was studied by Falconer, Gauld and Roberts (1978*b*). In all four organs the cells of the Large strain were larger than those of the Small at 6 weeks of age. The percentage difference is given in the right-hand column of Table 14. The correspondence for each organ is not close, but the averages of the four organs are not very different, 44 % for the localized effect in the chimaeras and 36 % for the cell size effect. The correspondence may be no more than a coincidence. We point it out to show that the localized effect of the cellular genotype on organ weight may be mediated through its effect on cell size.

(iv) *Chimaeric heterosis*

Chimaeras are generally acknowledged to be relatively large and healthy individuals. This raises the question of whether they benefit, in a manner analogous to heterosis, from the mixture of cell populations of different origin. We shall therefore examine the evidence for 'heterosis' for body weight in our material. First, however, we must know whether the strains used to make the chimaeras show heterosis in the ordinary genetic sense when crossed. Crosses were therefore made and it is sufficient to say that they showed on average about 12 % heterosis for weight at 6 weeks, heterosis being defined as the difference between the F_1 and mid-parent means.

There are two independent ways in which heterosis in the chimaeras can be looked for: both are suggestive but unfortunately inconclusive. The first way is by consideration of the elevation of the regression line in Fig. 2. If there were no heterosis the points would be equally distributed above and below the broken line joining relative weights of 0 and 100. It is obvious that there are more points above than below. The observed regression does not differ from the broken line in slope but it does so in elevation. The predicted mean relative weight at $\bar{P}_w = 50$ is $59{\cdot}6 \pm 3{\cdot}3$, which is significantly greater than the value of 50 % expected with no heterosis ($t_{45} = 2{\cdot}91$; $P < 0{\cdot}01$). This therefore looks like convincing evidence for chimaeric heterosis. The significance test, however, takes no account of the errors in estimating the mean weights of the constituent strains, and it is not possible to arrive at a reliable figure for this error. It is noticeable that most of the evidence for heterosis in Fig. 2 comes from one chimaera type, the C/L (see Table 11). We therefore think that the evidence from Fig. 2 cannot be regarded as proving the existence of chimaeric heterosis. There is, moreover, a possible reason for the evidence from the C/L chimaeras being spurious. The weights analysed were those at 6 weeks but the cell proportions of these chimaeras were determined at 50 weeks, and the mean of $\bar{P}$ was well below 50 % (Table 5). If there were a progressive reduction of the proportion of 'large' cells with increasing

age, the body weights at 6 weeks would be above their expectations based on the cell proportions at 50 weeks, giving the appearance of heterosis.

The second way of looking for heterosis is more direct. It involves the comparison of the weights of overt chimaeras with those of single-component chimaeras. The overt chimaeras are those of types L/L, C/C and S/S. Here the constituent strains are of similar size and the variation of cell proportions does not have much effect on weight. The single component chimaeras are from all types; so that, for example, single-component individuals with 100% of cells from one or other of the Large strains are compared with overt chimaeras of the type L/L. Similarly, single-component chimaeras with C cells are compared with C/C overts; and single-component S with S/S. The results are given in Table 15. In all three size-types the overt chimaeras are heavier than the single-component chimaeras, but

Table 15. *Comparisons of the weights of overt chimaeras of like-sized strains with those of single-component chimaeras*

	Size		
Overt	Large	Control	Small
Source	L/L	C/C	S/S
Number	3(2)	5	8
Mean weight (g)	39·47 (35·70)	26·99	24·78
Single-component			
Source	L/S, S/L, C/L, L/L	L/C, C/L, C/C	S/L, S/S
Number	11	10	3
Mean weight (g)	34·44	26·35	21·20

not by much. To assess the significance of the difference, the data were subjected to a two-way analysis of variance, treating overt *vs.* single-component as a fixed factor, weights being first transformed to logs. This gave an F-ratio of 5·7 for 1 and 34 d.f., with P = 0·02. There is, however, a difficulty in accepting this as conclusive evidence of heterosis. One overt L/L individual had an exceptionally high weight of 47·0 g. The mean with this individual omitted is shown in parentheses in Table 15. The analysis of variance with this individual omitted gave $F = 3{\cdot}2$, P = 0·08, which is not significant. So the evidence for heterosis rests heavily on a single individual and cannot be accepted with confidence.

To summarize: the two independent lines of evidence both suggest that chimaeras show 'heterosis' for body weight, but both comparisons suffer defects which make the conclusion not completely convincing.

5. DISCUSSION

The results have shown clearly that none of the organs studied plays a predominant role in controlling growth. With the possible exception of the pituitary, this is not really surprising. Nevertheless, all appear to play some part in controlling growth because body weight is linearly dependent on the mean cell

proportions in these eleven organs. Can one conclude from this that growth is controlled by the cellular genotype throughout the body; or is there some other organ, not studied, that controls growth? Unfortunately this question cannot be answered conclusively. Evidence for the existence of such an organ would come from chimaeric variance not accounted for by the mean of the organs studied. The estimate of this variance, 4·2 g^2, was not significantly different from zero. All that can be said, therefore, is that we have no compelling evidence for the existence of a growth-controlling organ other than those studied. The absence of a growth-controlling organ is suggested by a study of embryonic growth (Gauld, 1980). From 11 days of gestation till birth embryos of the large strains were found to be heavier than embryos of the small strains, and the differences could not all be attributed to maternal effects. At 11 days, organogenesis has barely started and no organ has completed its differentiation. The embryonic difference in weight can therefore hardly be attributed to any specific organ or tissue. The difference of embryonic weights, after subtracting the estimated maternal effect, amounted to about 11 percent. The much larger difference developed postnatally could be due to a different growth-controlling mechanism.

The mean cell proportions, over all chimaeras made from strains of different sizes, was close to 50%. From this we drew the conclusion that there was no differential cell proliferation: cells from the larger component did not tend to outgrow those from the smaller. One must, however, ask what cell proportions would be expected if the cells proliferated at the rates characteristic of their strains of origin; would it be detectably different from 50%? Mice of the large strains have more cells than those of the small strains at the same age of 6 weeks. The cells of the large strains must therefore proliferate faster than those of the small. Falconer, Gauld & Roberts (1978*b*) give estimates of the total cell numbers in four organs (lung, liver, spleen and kidney) of Large, Control and Small strains. Let N_L and N_S be the cell numbers in an organ of the larger and smaller strains used to make a chimaera. Suppose that the organ in a chimaera starts with 50% of 'large' cells and that the cells subsequently proliferate at their own intrinsic rate. The adult organ will then contain $\frac{1}{2}(N_L+N_S)$ cells. Provided the mean initial proportion is 50 percent, the mean proportion of 'large' cells in the adult organs will then be $N_L/(N_L+N_S)$. Taking the values of N_L, N_S and N_C (for controls) from Table 3 of Falconer, Gauld & Roberts (1978*b*) allows us to calculate the expected cell proportions, $E(\bar{P})$, for the chimaeras of each type, assuming that the cell numbers in the organs studied here are the same on average as the mean of the four organs for which N is known. The expected cell proportions are then as follows:

Chimaera types	L/S, S/L	L/C, C/L	S/C
$E(\bar{P})$(%)	61·8	54·3	57·5

The observed values of $\bar{P}$ (Table 5) differ significantly from these expectations in both C/L and S/L (t = 2·5) and in the weighted mean of all unlike-size chimaeras (t = 4·2). We can therefore conclude that the rate of cell proliferation is not cell-specific in the chimaeras.

There are three levels at which the control of growth might be exercised: (1) systemic, all organs being subject to the same control; if there is a single growth controlling organ, it would have to operate in this way, (2) at the level of organs, each organ having its growth determined by its own cellular composition, and (3) at the cellular level, the rate of proliferation being cell-specific. The last of these possibilities is disproved by the consideration of mean cell proportions in the previous paragraph. The second possibility was disproved as the main way by which growth is controlled. It was shown that the weights of some organs are probably influenced by their own cellular composition, but this effect was not nearly enough to account for the differences in body weight associated with mean cell proportions. We are therefore left with the conclusion that growth control must be mainly systemic. But whether the control originates from a particular organ, not among those studied, or is in some undefined way dependent on the overall cellular composition of the body cannot be inferred from the present results.

One aspect of the cell proportions in the organs remains to be discussed, and that is the puzzling correlation, found for some organs, between $P-\bar{P}_w$ and $\bar{P}_w$; i.e. between the deviation of the organ from the mean of the individual and the mean itself. In particular, this correlation was negative for the brain and positive for the blood. The following two developmental models, though not very plausible, may be suggested as ways by which these correlations could arise. Both require the supposition that there is a tendency for the cell proportions to change during development and to change, moreover, in the direction of one or other extreme. Such a change might result from the majority cell-type inhibiting the proliferation of the minority type. In one model the change toward the extremes takes place in the undifferentiated tissues, from which the organs become differentiated sequentially. After differentiation the organs do not change further in their cell proportions. Thus the first-formed organs will have their cell proportions closer to the initial value and the later-formed organs will have them further toward the extremes. This would generate a negative correlation of $P-\bar{P}_w$ with $\bar{P}_w$ in the first-formed organs and a positive correlation in the later-formed ones. In the second model, which is perhaps somewhat less implausible, the change toward the extremes takes place in all organs during the whole course of their development, and the organs do not need to differentiate sequentially. Organs with little cell replacement would change least and would have a negative correlation of $P-\bar{P}_w$ with $\bar{P}_w$; organs with much cell replacement would change most and have a positive correlation. The correlations observed for the brain and blood fit with this expectation. The spinal cord, which would be expected to be like the brain, has also a negative correlation though a smaller one. If this second model were right, one might expect to find changes of cell proportions during the life of the individual, and this can be tested from the blood which was assayed at two ages. The expectation from the model is that the variance would increase with age. The blood was assayed at 6 and 50 weeks in the C/L chimaeras and at 4 and 6 weeks in the others. The variance increased in the C/L, but not significantly ($P \sim 0{\cdot}2$) and it decreased non-significantly in the others, so this test gives no support for this

model. We must therefore leave these correlations as an unexplained feature of the results.

We are indebted to Dr Patricia Bowman who made some of the chimaeras.

REFERENCES

BOWMAN, P. & MCLAREN, A. (1970). Viability and growth of mouse embryos after *in vitro* culture and fusion. *Journal of Embryology and Experimental Morphology* **23**, 693–704.

FALCONER, D. S. (1973). Replicated selection for body weight in mice. *Genetical Research* **22**, 291–321.

FALCONER, D. S. & AVERY, P. (1978). Variability of chimaeras and mosaics. *Journal of Embryology and Experimental Morphology* **43**, 195–219.

FALCONER, D. S., GAULD, I. K. & ROBERTS, R. C. (1978*a*). Growth control in chimaeras. In *Genetic Mosaics and Chimaeras in Mammals* (ed. L. B. Russell), pp. 39–49. New York: Plenum Press.

FALCONER, D. S., GAULD, I. K. & ROBERTS, R. C. (1978*b*). Cell numbers and cell sizes in organs of mice selected for large and small body size. *Genetical Research* **31**, 287–301.

GARNETT, I. & FALCONER, D. S. (1975). Protein variation in strains of mice differing in body size. *Genetical Research* **25**, 45–57.

GAULD, I. K. (1980). Prenatal growth and development in fast and slow growing strains of mice. Ph.D. Thesis, University of Edinburgh.

KLEBE, R. J. (1975). A simple method for the quantitation of isozyme patterns. *Biochemical Genetics* **13**, 805–812.

MCLAREN, A. (1976). *Mammalian Chimaeras*. Cambridge University Press.

MOORE, W. J. & MINTZ, B. (1972). Clonal model of vertebral column and skull development derived from genetically mosaic skeletons in allophenic mice. *Developmental Biology* **27**, 55–70.

NESBITT, M. N. (1978). Attempts at locating the site of action of genes affecting behaviour. In *Genetic Mosaics and Chimaeras in Mammals* (ed. L. B. Russell), pp. 51–58. New York: Plenum Press.

PADUA, R. A., BULFIELD, G. & PETERS, J. (1978). Biochemical genetics of a new glucose-phosphate isomerase allele (*Gpi-1*^{c}) from wild mice. *Biochemical Genetics* **16**, 127–143.

ROBERTS, R. C., FALCONER, D. S., BOWMAN, P., & GAULD, I. K. (1976). Growth regulation in chimaeras between large and small mice. *Nature* (*Lond*) **260**, 244–245.

SHAW, C. R. & PRASAD, R. (1970). Starch gel electrophoresis of enzymes – A compilation of recipes. *Biochemical Genetics*, **4**, 297–320.

PAPER 18

Cell numbers and cell sizes in organs of mice selected for large and small body size.

Genet. Res. 31, 287-301. 1978

by

D.S. FALCONER, I.K. GAULD & R.C. ROBERTS

Genet. Res., Camb. (1978), **31**, *pp.* 287–301
With 5 text-figures
Printed in Great Britain

Cell numbers and cell sizes in organs of mice selected for large and small body size

By D. S. FALCONER, I. K. GAULD and R. C. ROBERTS
*Agricultural Research Council, Unit of Animal Genetics**

(*Received* 10 *March* 1978)

SUMMARY

Cell numbers in four organs of large, control and small mice were estimated by nuclear counts. Average cell mass was estimated from the cell number and the organ weight. The mice were from the selected Q-strain with six replicate lines in each size-group. The organs were lung, liver, spleen and kidney. At 6 weeks of age the large mice had more cells and larger cells than the controls in all organs; the small mice had fewer and smaller cells than the controls. The regression of log cell-number on log-organ weight provides a measure of how much, proportionately, cell number contributes to the differences in organ weight. In the lung and spleen, cell number contributed about 70% of the strain differences in organ weight, cell mass contributing about 30%; in the liver and kidney the relative contributions were about equal, at 50%.

Cell counts at different ages from 3 to 15 weeks showed that cell number and cell mass contributed to the increases of organ weights during growth in roughly the same proportions as stated above. From this it is concluded that the main effect of selection for body weight has been to speed up or slow down the normal processes of cellular growth.

1. INTRODUCTION

When body size is changed by selection, the response might be partitioned into changes in the number of cells and in average cell size. These changes could then be described formally in terms of the genetic correlations of cell number and cell mass with body weight. An alternative viewpoint is to consider a genetic change in body weight as an adjustment of the regulation of growth, and to ask whether this regulation operates by changing cell number or by changing cell size, or by both.

The relation of body size to cell number and cell size in *Drosophila* has been very thoroughly studied by Robertson (1959*a*, *b*). Genetic variation of both number and size was found, with the interesting difference that the genetic variation of cell number was mainly additive, but that of cell size was non-additive. Information about mammals is much less complete, though genetic variation in both cell number and cell size have been reported. Robinson & Bradford (1969) found that the larger sizes of seven organs in a strain of mice selected for increased post-weaning growth were due to increased cell number. Hanrahan, Hooper & McCarthy (1973)

* Address: Institute of Animal Genetics, West Mains Road, Edinburgh EH9 3JN, Scotland.

studied muscles in strains of mice selected up and down for growth rate, and found the changes in muscle size were mainly due to fibre number though there were some changes also in fibre diameter. Krzanowska (1967) compared F_1 hybrid mice with the parental inbreds and found that the heterosis in the growth of the embryonic liver was due to cell number. Musialek (1974) similarly compared F_1 and inbred mice and found postnatal heterosis due mainly to cell number, with smaller effects due to cell size. Priestley & Robertson (1973), studying the same strains as are described in this paper, though by different methods, found the main differences to be in cell number, but there were smaller differences in cell size. The general conclusion from the previous work on mice is that genetic differences of size are mainly, but not exclusively, due to cell number.

There have been many studies of environmental effects on cellular properties. The general conclusion from the work on nutrition in rats is that the level of nutrition early in postnatal life affects cell number and not cell size, but later in life it affects cell size and not cell number (Winick & Noble, 1966, 1967; Winick, Fish & Rosso, 1968). Musialek (1974) compared the effects of nutritional level with those of heterosis and found both affected cell number and cell size in the same ways.

Differences of size between mammalian species are mainly differences of cell number (Berrill, 1955), so one might expect artificial selection to affect mainly cell number. Cell size does, however, differ between species; the cells of mice and elephants differ by a factor of 2 in linear dimensions, and so by a factor of 8 in volume (Berrill, 1955). Furthermore, cell size increases during the growth of the individual, a 4-fold increase occurring in the rat kidneys after birth (Winick & Noble, 1965).

From the evidence of previous work it seems, therefore, that one should expect there to be genetic variation of both cell number and cell size on which artificial selection for body weight might act. This paper examines the differences of cell number and cell size between strains of mice previously selected up and down for body weight. The weight of the large mice was about twice that of the small at 6 weeks of age. An important aspect of the material was that the selection was replicated. There were six lines selected independently for large size, six selected for small size and six unselected controls. Any change that is found regularly in all the replicates can with more confidence be ascribed to the genetic differences in growth rate, whereas irregular changes, differing from line to line, are more likely to be the consequences of random drift, perhaps unrelated to the character selected for.

2. MATERIALS AND METHODS

(i) *Sources of mice*

Two sets of data were obtained. The first, or main, experiment was a 'cross-sectional' study in which the material was obtained from mice all aged 6 weeks. The second, or subsidiary, experiment was a 'longitudinal' study in which material was obtained from a smaller number of mice at six different ages. All the mice

came from the replicated Q-lines selected for body weight at 6 weeks of age (Falconer, 1973). There were six Large (L) lines, six Small (S) lines and six unselected Control (C) lines. The six replicates within each size-group (i.e. L, C, or S) were labelled A–F, so that, for example, LA and SA are the large and small lines of the A-replicate. The three lines (L, C and S) of each replicate shared some common ancestry in the base population, but no resulting correlations in any feature were found. The 18 lines are therefore best regarded as 6 random replicates in each of the three size-groups.

The mice for the main experiment came from the 14th and 15th generations, when the lines had made about 85% of the total response achieved by generation 21. After generation 21 all the lines were maintained without selection, and the mice for the subsidiary experiment came from generation 31. For the main experiment each of the 18 lines provided 8 males and 8 females taken equally from two litters in each of the two generations. There were thus in all 144 mice of each sex. The subsidiary experiment was restricted to males of only two replicates, B and E, in each size-group, making 6 lines in all. Material was obtained from mice at six ages, namely 3, 5, 6, 7, 9 and 15 weeks. Each line provided 4 mice, all males, at each age, taken from two litters. Some litters contributed to several ages, others to only one. The total number of mice was 144.

(ii) *Cell-counting*

The procedure for preparation of the tissues and counting of the cells was the same in both experiments. After the mice were killed, they were bled thoroughly. Four organs – lung, liver, spleen and kidneys – were weighed, and homogenized at constant speed for exactly 1 min in 25 ml (50 ml for liver) of 0·01 N-HCl. The method was basically that described by Zumoff & Pachter (1964) for releasing nuclei for counting. The homogenate was examined microscopically for residual clumping of cells. In a few cases, further homogenization was carried out, but the practice was avoided in marginal cases to reduce the risk of breaking nuclei. The homogenate was stored at 4 °C to await counting; the storage period did not affect mean nuclear counts. Care was taken that the size-groups did not differ much in their mean duration of storage, or in the range of dates over which their cells were counted. Five samples were taken from each aliquot, and each sample was counted on a haemocytometer slide. The nuclei were counted in five areas spaced systematically on each slide. The total volume in which nuclei were counted for each organ was 1×10^{-4} ml. Multiplying the total count by 25×10^4 (or by 50×10^4 for liver) gave the estimate of the cell-number in the organ. An estimate of cell size was obtained by dividing the organ weight by the number of cells in that organ. All extracellular components will of course affect the estimate, though there is no particular reason for this error to affect the size-groups differentially. The measure of cell size is thus the weight of organ associated with each nucleus, for which we use the term 'cell mass', expressed in nanograms ($g \times 10^{-9}$).

(iii) *Error variance*

The design of the experiment provided estimates of the error variance of the counts of nuclei, which could be partitioned into components between squares within slides, and between slides. It did not, however, allow the whole of the error variance between individual mice to be estimated, because each organ was homogenized as a whole. After the data had been collected it appeared that there were some differences between mice that were far too great to be real. These differences must have arisen, in part, from 'error' in the preparation of the suspension of nuclei for counting, but we have no means of estimating this error variance, and so we cannot assess the significance of differences between individual mice. The between-mice component is used as the error for assessing differences between lines.

For each organ we have three parameters whose interrelations are to be studied: organ weight, cell number and cell mass. It is important to note, however, that we have only two independent variables: organ weight and cell number, cell mass being derived directly from these two. There is no reason to suppose that the error deviations in organ weight and cell number will be correlated, but the error deviations in cell number and cell mass are correlated negatively. For this reason we can get no information about any real correlation that there may be between cell number and cell mass.

Table 1

(*a*) Six-week body weight (g) of the mice used in the main study, 16 mice per line, sexes averaged

	A	B	C	D	E	F	Mean
L	32·86	32·76	32·06	31·57	30·54	34·26	32·34
C	25·11	28·42	22·39	24·80	23·36	22·78	24·48
S	16·12	17·56	17·22	15·99	17·00	15;14	16·51

(*b*) Mean weights (g) of the organs of the mice used, pooled over replicates

	Lung	Liver	Spleen	Kidney
L	0·221	2·438	0·136	0·549
C	0·167	1·683	0·103	0·405
S	0·122	1·038	0·058	0·252

3. RESULTS

The results of the main experiment will be presented first; those of the longitudinal study are presented in section (v) below.

(i) *Body weight and organ weight*

The mean body weights of all the lines at 6 weeks are given in Table 1(*a*). Except for a rather high value in the CB line, the samples are representative of the lines from which they were drawn, as described by Falconer (1973). Table 1(*b*) gives the mean organ weights in the three size-groups.

In order to see how the organ weights were related to body weight, the mean log organ-weight of each line was plotted against the mean log body-weight,

plotting each sex separately but on the same graph. These plots, shown in Fig. 1, are all clearly linear, justifying the calculation of linear regressions. In no organ was the slope of the regression line significantly different between the sexes. Common regressions pooled within sexes were therefore calculated and these are

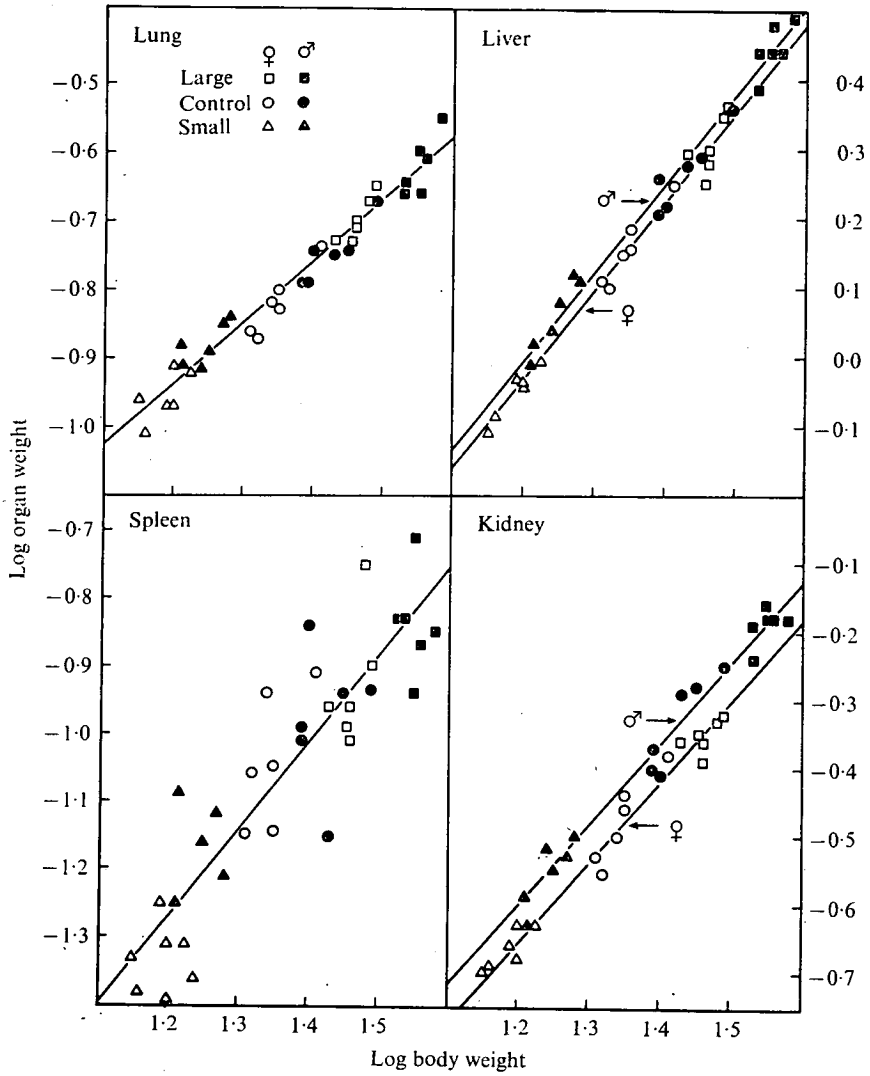


Fig. 1. Relation of organ weight to body weight, with regression of $\log_{10}$ organ-weight on $\log_{10}$ body-weight, in mice aged 6 weeks. The scales for liver and kidney are at the right.

shown in the figure. The elevations of the regression lines of the two sexes did not differ significantly in lung and spleen, and single lines are drawn for these organs. In the liver and kidney, however, the sexes differed significantly in elevation ($P < 0{\cdot}001$ in both cases), males having relatively larger organs than females. Males had livers 6·1 % heavier than females and kidneys 13·3 % heavier.

For each organ the regression equation $\log y = \log a + b(\log x)$ was calculated, y being organ weight and x body weight. This gave the allometric relation $y = ax^b$. The values of a and b are given in Table 2. All the estimates of the common b are significantly different from 1, being less than 1 for lungs and greater than 1 for the other organs. In other words, large mice have relatively smaller lungs, and relatively larger livers, spleens and kidneys.

Table 2. *Relation between organ weight (y) in grams and body weight (x), in grams, from $y = ax^b$, in mice aged 6 weeks*

(a and b were estimated from regressions of log y on log x. The values of b are the common regression coefficients within sexes. The values of a are derived from the common b and the separate means of each sex.)

	$a_{♀}$	$a_{♂}$	$b \pm$ S.E.
Lung	0·0098	0·0100	0·888 ± 0·033
Liver	0·0273	0·0290	1·277 ± 0·029
Spleen	0·0016	0·0016	1·274 ± 0·120
Kidney	0·0087	0·0098	1·173 ± 0·032

(ii) *Cell size and cell mass*

The changes in cell size and cell mass that have resulted from selection for body weight are first illustrated diagrammatically in Fig. 2. The three line-means (L, C, S) in each replication (A–F) are connected as if they were one-step correlated responses in six separate selection experiments, with the Control line as the starting point. The changes of cell number are very consistent, the order of the lines being L > C > S in all four organs and all replicates, except for the liver in two replicates. The changes in cell mass are also consistent in showing that downward selection has decreased cell mass, i.e. C > S, to which there are two exceptions; upward selection was less consistent, with five exceptions to the order L > C. The presentation in Fig. 2 thus leaves no doubt that both cell number and cell mass have been changed in all the organs. As noted earlier, however, the significance of the changes should be assessed by treating replicates as random lines within size-groups. This was done as follows.

The means of the size-groups were calculated from the six line-means in each, and the differences tested by t-tests. The results are given in Table 3. These again leave no doubt that the selection for body weight has changed both cell number and cell mass in all of the four organs. The proportionate changes are given in Table 4, with the changes in organ weight for comparison. The proportionate changes are rather more in cell number than in cell mass, particularly in lung and spleen. We shall return in the next section to the question of how much of the changes of organ weight are attributable to cell number and how much to cell mass.

Hierarchical analyses of variance were also carried out to see whether there were significant differences between lines within size-groups. The analyses of variance are not given in full, but only a summary of the components, in Table 5. The components of cell number and cell mass between replicate lines within size-groups

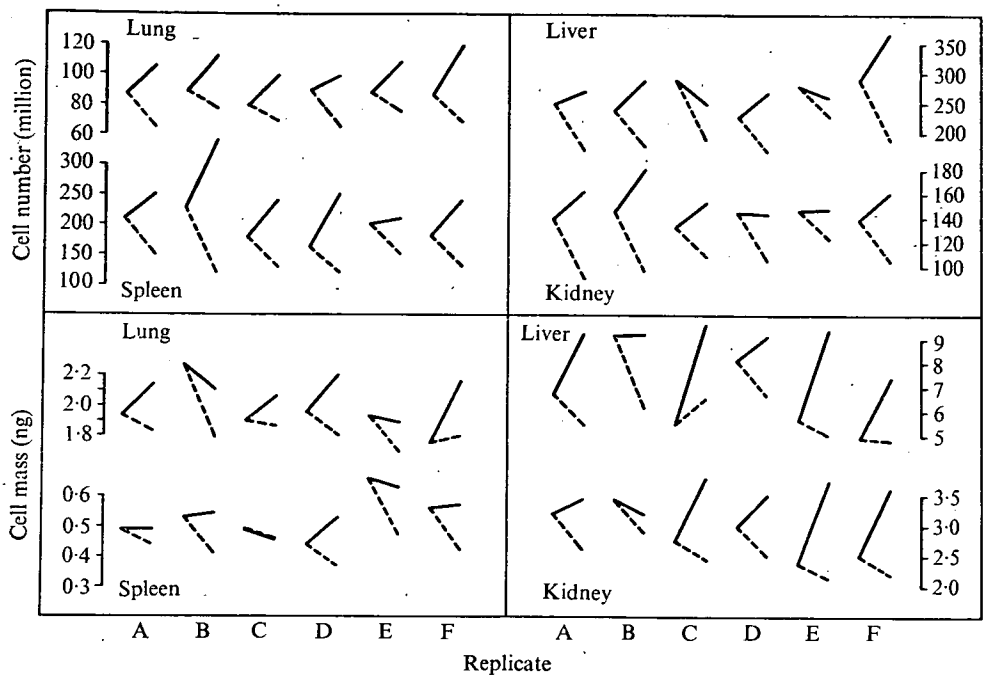


Fig. 2. Changes in cell number and cell mass in mice aged 6 weeks brought about by selection for body weight. The six replicates are depicted as separate 'one-step' selection responses, the control levels being taken as the starting points. Solid lines represent the responses to selection for large body size, broken lines selection for small body size. The sexes are averaged. The scales for liver and kidney are at the right.

Table 3. *Mean cell number and cell mass, with standard errors, in the three size-groups, Large (L), Control (C) and Small (S)*

(Each mean is based on six line-means, sexes averaged. The stars give the significance of the differences between size-groups.)

	Lung	Liver	Spleen	Kidney
		Number (millions)		
L	107·1 ± 3.2	286·6 ± 18·1	257·5 ± 19·1	158·6 ± 5·5
L–C	**		*	*
C	86·1 ± 1·5	264·8 ± 10·5	193·9 ± 10·6	142·0 ± 2·1
C–S	***	***	***	***
S	69·1 ± 2·3	191·5 ± 8·3	132·9 ± 5·7	104·2 ± 4·6
L–S	***	***	***	***
		Mass (ng)		
L	2·094 ± 0·044	9·018 ± 0·336	0·537 ± 0·024	3·548 ± 0·086
L–C		*		**
C	1·956 ± 0·069	6·681 ± 0·678	0·530 ± 0·031	2·895 ± 0·162
C–S	*		*	
S	1·792 ± 0·026	5·739 ± 0·286	0·432 ± 0·015	2·464 ± 0·112
L–S	***	***	**	***

* $P < 0·05$, ** $P < 0·01$ *** $P < 0·001$

were all significant at the 5% or higher level in both sexes in all organs, with the exception only of cell number in female lungs.

The component between size-groups was greater than the component between replicates within size-groups in all cases, but the difference was much greater for cell number than for cell mass. These comparisons are given at the foot of Table 5.

Table 4. *Proportionate changes in organ weight, cell number and cell mass, based on the size-group means in Table 3*

(Each entry is the percentage difference from Control.)

		Lung	Liver	Spleen	Kidney	Mean
Organ weight	L–C	33	46	34	38	38
	S–C	−26	−38	−44	−38	−36
Cell number	L–C	24	8	33	12	19
	S–C	−20	−28	−32	−27	−27
Cell mass	L–C	7	35	1	23	16
	S–C	−8	−14	−18	−15	−14

Table 5. *Components of variance of cell number and cell mass in mice aged 6 weeks, sexes averaged*

(The total variance given is the sum of the components in actual units. The components are given in percentages of the total. The components between replicates are within size-groups, and those of individuals are within replicates. For explanation of the ratio of components, see text.)

	Lung	Liver	Spleen	Kidney	Mean
Cell number					
Total (millions)^2	534	6426	6420	1394	—
Size-groups (%)	67	37	59	55	54
Replicates (%)	5	10	14	5	9
Individuals (%)	28	53	27	40	37
Cell mass					
Total (ng)^2	0·1093	7·735	0·0157	0·7364	—
Size-groups (%)	20	34	18	38	27
Replicates (%)	10	11	16	10	12
Individuals (%)	70	55	66	52	61
Ratio of components					
Size-groups/replicates					
Cell number	13·4	3·7	4·2	11·0	6·0
Cell mass	2·0	3·1	1·1	3·8	2·25

In the lung, for example, the ratio of the component between size-groups to the component between replicates is 13·4 for cell number but only 2·0 for cell mass. The differences are in the same direction in the other organs, though quite small in the liver. The differences between size-groups were the result of selection while the differences between replicates were mainly the result of random drift. It seems, therefore, that the genetic changes brought about by selection have affected cell number relatively more than have the genetic changes resulting from random drift.

(iii) *Relative importance of cell number and cell mass*

The proportionate changes described in the previous section suggest that the differences of organ weights have been brought about on the whole more by changes in cell number than by changes in cell mass. The relative contribution that each has made to the differences of organ weight can be quantified from the regression of log cell-number on log organ-weight, as the following considerations will show.

If the logarithms of cell number, cell mass, and organ weight are denoted by n, m, and w respectively, then

$$n+m = w,$$

$$\text{cov}_{nw} = \text{cov}_{(w-m)w} = \text{var}_w - \text{cov}_{mw}$$

Dividing both sides by var_w gives

$$b_{nw} = 1-b_{mw},$$

$$b_{nw}+b_{mw} = 1,$$

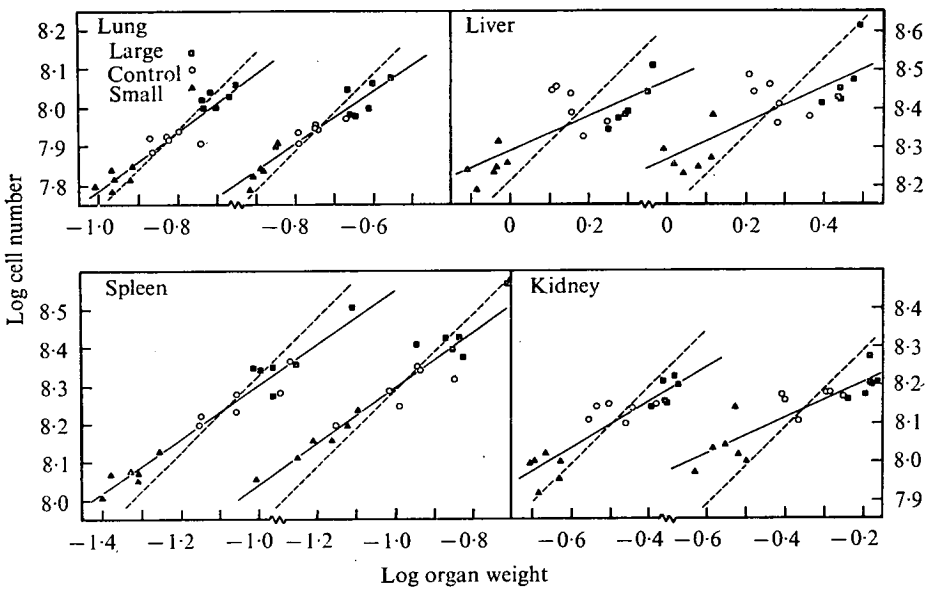


Fig. 3. Regressions of $\log_{10}$ cell-number on $\log_{10}$ organ-weight in mice aged 6 weeks. The continuous straight lines are the calculated regressions. The broken lines have a slope of 1, as expected if the differences of cell number explained all the differences of organ weight. Females are on the left, males on the right, in each organ.

where b_{nw} and b_{mw} are the regression coefficients of log cell-number and of log cell-mass respectively on log organ-weight. As noted earlier, there was error of unknown amount in estimating cell number. The error in estimating organ weight, however, was negligible, so the estimation of b_{nw} is valid.

The regression n on w therefore provides a measure of the relative contribution of cell number to the differences of organ weight, ranging from $b_{nw} = 0$, when the

whole difference is due to cell mass, to $b_{nw} = 1$, when the whole difference is due to cell number.

Plots of log cell-number on log organ-weight are shown in Fig. 3. Linear regressions were calculated from the line-means separately for each sex, and are shown by solid lines in the figure. The broken lines have slopes of 1, showing where the regressions would lie if cell number were wholly responsible for the differences of organ weight. The corresponding numerical values are given in Table 6A. The intercept, log a, is the predicted log cell-number of an organ weighing 1 g, and it provides a measure of the elevation of the regression line. The regression coefficients, b, are all less than 1. All four regressions in each sex are significantly different from both 1 and zero with $P < 0{\cdot}001$ in every case. The regressions in the two sexes are not significantly different in any organ and are combined in the common regression given in Table 6A. These show that changes in cell number

Table 6. *Relations of cell number (N) to organ weight (W) in grams by the regression $\log N = \log a + b \log W$, calculated from line-means, with standard errors*

(A: All mice aged 6 weeks. B: Mice aged 3–15 weeks (3–6 weeks for liver). The values of a for the two sexes at 6 weeks are calculated from the common regression.)

	Lung	Liver	Spleen	Kidney
	(A) Age 6 weeks			
Females				
Log a	8·544 ± 0·049	8·291 ± 0·019	9·004 ± 0·046	8·400 ± 0·035
Log b	0·757 ± 0·059	0·453 ± 0·090	0·704 ± 0·041	0·603 ± 0·068
Males				
Log a	8·452 ± 0·041	8·270 ± 0·028	9·018 ± 0·057	8·298 ± 0·025
Log b	0·682 ± 0·054	0·470 ± 0·091	0·723 ± 0·056	0·460 ± 0·064
Common b	0·719 ± 0·040	0·462 ± 0·063	0·712 ± 0·034	0·521 ± 0·048
♀ log a	8·512	8·289	9·014	8·359
♂ log a	8·480	8·272	9·007	8·320
N (♀)/N (♂) = M (♂)/M (♀)	1·076	1·040	1·016	1·094
	(B) Age 3–15 weeks*			
Males				
Log a	8·490 ± 0·052	8·204 ± 0·025	9·157 ± 0·058	8·405 ± 0·019
Log b	0·603 ± 0·059	0·269 ± 0·090	0·868 ± 0·053	0·552 ± 0·037

* 3–6 weeks for the liver.

account for about 70 % of the differences of organ weight in lung and spleen, and for about 50 % in liver and kidney. (Lung and spleen are not significantly different from each other, and nor are liver and kidney; but both lung and spleen are significantly different from both liver and kidney, with $P < 0{\cdot}01$, or $P < 0{\cdot}001$.) Complementarily, the relative contribution of cell mass to the differences of organ weight between lines, measured as $1 - b_{nw}$, was about 30 % for lung and spleen and about 50 % for liver and kidney. These estimates confirm and quantify the impression given by the simple treatment in the previous section.

(iv) *Comparison of sexes*

Males have larger organs than females. How much of this difference is due to cell number and how much to cell mass? As noted earlier, the regressions of log cell-number on log organ-weight in males and in females did not differ in slope. The common regression was calculated and the two regression lines were tested for differences in elevation to make the comparison of log cell-number at the same organ weight. The elevations were not significantly different in liver or in spleen, but they were in lung ($P < 0{\cdot}01$) and in kidney ($P < 0{\cdot}05$). In these organs males had fewer and larger cells than females. As a measure of elevation the intercepts were calculated from the common regression. The antilog of the difference between the intercepts gives the cell number in one sex relative to that in the other when adjusted to the same organ weight. The cell number in females relative to males is the same as the cell mass in males relative to females. These relative values are given in Table 6A. Expressed in terms of cell mass, males had larger cells than females in all organs; in the lung they were 7·6 % larger and in the kidney 9·4 % larger. The differences of 4·0 % in the liver and 1·6 % in the spleen were not significant, as noted earlier.

To estimate the relative contribution of cell number and cell mass to the sex-difference in organ weight, we need the regression based on the sex-means, i.e. the between-sex regression. This was 0·30 in lungs and 0·25 in kidneys, so the difference between the sexes in the weights of these organs was 70 and 75 % due to cell mass, in contrast to 22 and 48 % for the differences between the lines.

(v) *Changes during growth*

The main experiment has shown that the Large, Control and Small mice differed in cell mass in the four organs, when compared at the fixed age of 6 weeks. Data for the longitudinal study, to be described now, were collected with the object of finding out if the cellular changes during growth resembled those brought about by selection. Cell mass is known to increase during growth in several organs and tissues of rats (Enesco & Lablond, 1962; Winick & Noble, 1965). If the same is true of the organs studied in our mice, selection could have produced the observed differences of cell mass by speeding up or slowing down this normal increase of cell mass during growth.

Fig. 4 shows the changes of cell number and cell mass during growth from 3 to 15 weeks. The two replicates in each size-group have been averaged since the mean of each line at each age was based on only four animals. Many irregularities remain in the graphs, but three features seem clear, if some exceptions are disregarded. (1) The size-groups differ in cell number in the expected direction at all ages. (2) Cell number increases from 3 to 6 or 7 weeks and then remains constant, or declines. It is hard to understand the decline of cell number in the lung; in the spleen it was accompanied by a reduction in organ weight; in the liver, where it is most marked, it could be due to the formation of polyploid cells. (3) Cell mass increases fairly regularly in all organs throughout the period from 3 to 15 weeks.

Except for the reduction of cell numbers, these changes during growth resemble in general outline those found by Enesco & Lablond (1962) and by Winick & Noble (1963) in rats, organ growth being mainly by cell number initially and by cell mass later. The increase of cell mass during the growth of the organs shows that the differences found at 6 weeks could be simply the developmental consequences of

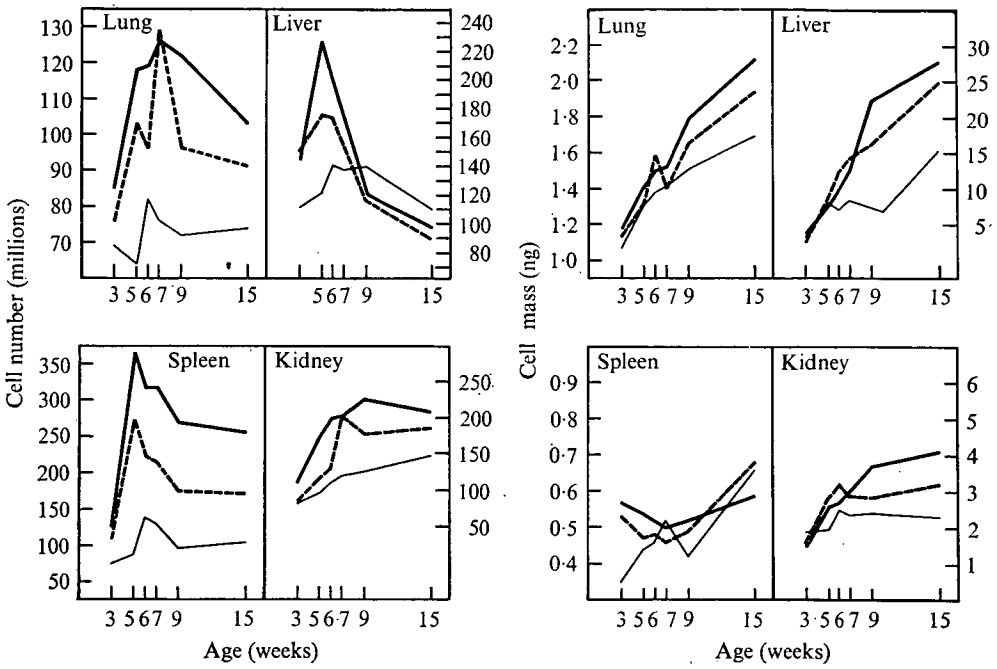


Fig. 4. Changes of cell number and cell mass in males during growth from 3 to 15 weeks. Means of two replicates each of Large (thick lines), Control (broken lines) and Small (thin lines).

the changes of organ weight brought about by selection. To test this possibility we analysed the data by regression in the manner described for the main experiment.

Fig. 5 shows the plots of log cell-number against log organ-weight, the points being line-means at each age. The essential difference between these graphs and those in Fig. 3 is that in Fig. 3 the differences of organ weight are due to the selection-history of the lines, whereas here (Fig. 5) they are due also to age-differences. The regression lines fitted to the points are shown on the graphs and the regression coefficients are given in Table 6B. The graph of the liver is confusing because of the marked reduction of cell numbers after about 6 weeks despite continued increase of organ weight. Because of the obvious non-linearity at the higher ages, the calculation of the regression in the liver was based on the points for 3, 5 and 6 weeks only.

With the possible exception of the liver, two main features of the results are clear. First, the points for the three size-groups and all ages fall reasonably well on

the same lines, showing that, in the main, organs of the same weight have the same cell mass, irrespective of age or of size-group. In the main, therefore, selection has not changed cell mass except as a concomitant to the change of organ weight during growth. Secondly, the slopes of the regressions do not differ much from those obtained from mice all aged 6 weeks given in Table 6A. The difference between the two regressions in males is not significant in any organ. The similarity of the two regressions shows that cell number and cell mass make roughly the same relative contribution to the increase of organ weight during growth as they do to the differences produced by selection.

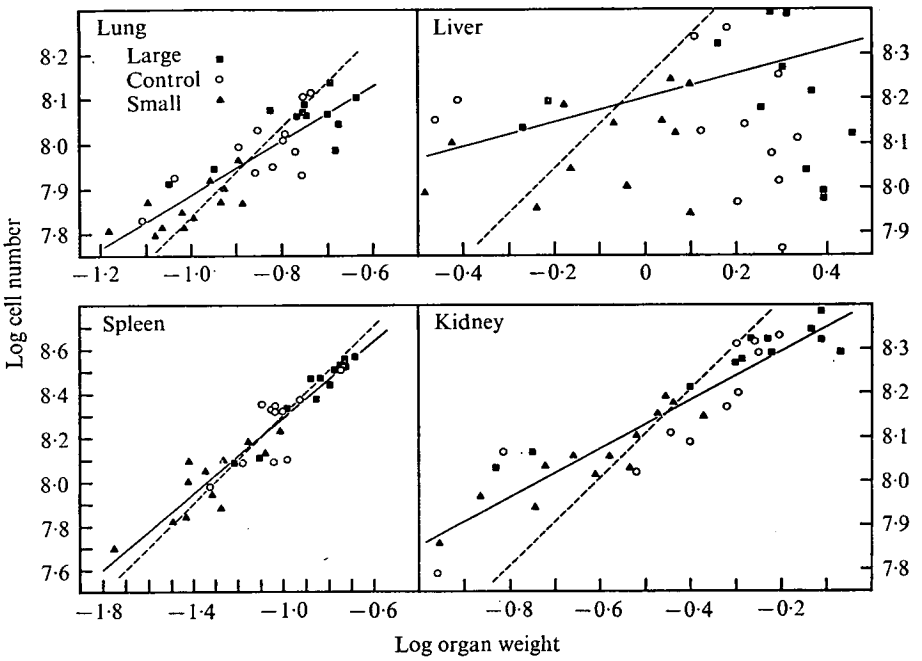


Fig. 5. Regressions of $\log_{10}$ cell-number on $\log_{10}$ organ-weight in male mice aged 3–15 weeks. The continuous straight lines are the calculated regressions (in the case of the liver, based on 3, 5 and 6 weeks only). The broken lines have a slope of 1, as expected if organs grew only by increase of cell number. Both scales in the graphs of spleen are half those of the other organs.

The conclusion to be drawn from the study of mice at different ages is that cell size increases during growth, and the difference in cell size between the selected lines is what would be expected from the different amounts of growth that they have made.

4. DISCUSSION

In the context of selection responses, the question asked was: did the response of body weight take place by changes of cell number or of cell size, and the answer was by both, in the four organs studied. But these two changes were themselves the consequence of a single effect of selection, the change in the rate of growth.

During the growth of any mouse the cells increase both in number and in size, the increase in cell size differing in amount between the organs. The effect of upward selection has been to make the mice grow faster so that at 6 weeks of age their cells are both more numerous and larger than those of the unselected controls. Downward selection had the opposite effect, resulting in mice at 6 weeks having fewer and smaller cells. When compared at the same body weight, and consequently at different ages, the Large, Control and Small strains had cells of roughly equal number and size in all the organs studied. The effect of selection might be summed up as a change in the relation of developmental age to chronological age.

The effects of selection for body size on the numbers and sizes of the cells of the lung, liver, spleen and kidney, described here, are the same as the effects on the numbers and diameters of muscle fibres reported by Byrne, Hooper & McCarthy (1973). The strains selected for increased and decreased growth rate, with which these authors worked, were derived from the same base population as the Q-stocks with which we worked. They measured the fibre number and diameter in seven muscles and found the large mice had consistently more and larger fibres than the controls, while the small mice had consistently fewer and smaller fibres. The muscle fibres of mice stop increasing in numbers soon after birth, and the subsequent increase of muscle size takes place by increase of the diameter of the fibres. Thus the developmental process in muscle fibres and in cells is similar in that both increase first in numbers and later in size. When the mice studied by Byrne, Hooper and McCarthy were compared at the same body weight the results were somewhat different from ours. At the same body weight, when the large mice were younger than the small, the large mice had more fibres than the small but with smaller diameters (Hooper & McCarthy, 1976). These results can be interpreted in the same way as ours: the large mice have gone through their developmental process faster than the small, but in this case fibre diameter increases with age independently of body weight. Consequently the younger large-strain mice have smaller fibres than the older small-strain mice.

A similar picture of the effect of selection on fatness was described by Clarke (1969). He studied the fat content of the same Q-strains after 14 generations of selection. The large mice had relatively more fat than the small at a fixed age, but when compared at the same weight there was little difference.

These three studies on the Q-strain mice show that selection for body weight has produced correlated responses in the numbers and sizes of cells in four organs, in the numbers and diameters of muscle fibres, and in the relative amount of fat. All these correlated responses have resulted from a single effect of selection in altering the timing of the normal developmental processes of growth.

We are greatly indebted to Dr St C. S. Taylor for helpful comments.

REFERENCES

BERRILL, N. J. (1955). The determination of size. In *Analysis of development* (ed. B. H. Willier, P. A. Weiss and V. Hamburger, pp. 620–630. London: Saunders.

BYRNE, I., HOOPER, J. C. & MCCARTHY, J. C. (1973). Effects of selection for body size on the weight and cellular structure of seven mouse muscles. *Animal Production*, **17**, 187–196.

CLARKE, J. N. (1969). Studies on the genetic control of growth in mice. Ph.D. thesis, University of Edinburgh.

ENESCO, M. & LEBLOND, C. P. (1962). Increase in cell number as a factor in the growth of the organs and tissues of the young male rat. *Journal of Embryology and Experimental Morphology* **10**, 530–562.

FALCONER, D. S. (1973). Replicated selection for body weight in mice. *Genetical Research* **22**, 291–321.

HANRAHAN, J. P., HOOPER, A. C. & MCCARTHY, J. C. (1973). Effects of divergent selection for body weight on fibre number and diameter in two mouse muscles. *Animal Production* **16**, 7–16.

HOOPER, A. C. B. & MCCARTHY, J. C. (1976). A note on fibre number and diameter in muscles of large and small lines of mice compared at a fixed body weight. *Animal Production* **22**, 131–133.

KRZANOWSKA, H. (1967). Cell number in relation to heterosis during embryonic growth of mice. *Genetical Research* **10**, 235–240.

MUSIALEK, B. (1974). The effect of genotype and environment on the size and number of cells in mouse organs. *Folia Biologica* **22**, 175–210.

PRIESTLEY, G. C. & ROBERTSON, M. S. M. (1973). Protein and nucleic acid metabolism in organs from mice selected for larger or smaller body size. *Genetical Research* **22**, 255–278.

ROBERTSON, F. W. (1959*a*). Studies in quantitative inheritance. XII. Cell size and number in relation to genetic and environmental variation of body size in Drosophila. *Genetics* **44**, 869–896.

ROBERTSON, F. W. (1959*b*). Studies in quantitative inheritance. XIII. Interrelations between genetic behaviour and development in the cellular constitution of the Drosophila wing. *Genetics* **44**, 1113–1130.

ROBINSON, D. W. & BRADFORD, G. E. (1969). Cellular response to selection for rapid growth in mice. *Growth* **33**, 221–229.

WINICK, M., FISH, J. & ROSSO, P. (1968). Cellular recovery in rat tissues after a brief period of neonatal malnutrition. *Journal of Nutrition* **95**, 623–626.

WINICK, M. & NOBLE, A. (1965). Quantitative changes in DNA, RNA and protein during prenatal and postnatal growth in the rat. *Developmental Biology* **12**, 451–466.

WINICK, M. & NOBLE, A. (1966). Cellular response in rats during malnutrition at various ages. *Journal of Nutrition* **89**, 300–306.

WINICK, M. & NOBLE, A. (1967). Cellular response with increased feeding in neonatal rats. *Journal of Nutrition* **91**, 179–182.

ZUMOFF, B. & PACHTER, M. R. (1964). Studies of rat kidney and liver growth using total nuclear counts. *American Journal of Anatomy* **114**, 479–493.

PAPER 19

Postnatal maternal effects on growth and fat deposition in mice selected for large and small size.

J. Anim. Sci. 53, 952-965. 1981

by

E.J. EISEN & R.C. ROBERTS

POSTNATAL MATERNAL EFFECTS ON GROWTH AND FAT DEPOSITION IN MICE SELECTED FOR LARGE AND SMALL SIZE[1]

E. J. Eisen[2,3] and R. C. Roberts[4]

Institute of Animal Genetics, West Mains Road, Edinburgh, Scotland EH9 3JN

Summary

A crossfostering experiment involving lines of mice selected for large (L) and small (S) 6-week body weight was designed to determine correlated responses in direct genetic and postnatal maternal genetic effects and postnatal litter size effects on fat deposition at 6 weeks of age. The gonadal fat pad was used as an index of adiposity. The L line exceeded the S line in both direct genetic and postnatal maternal genetic effects on weight and percentage (of body weight) of the gonadal fat pad. Postnatal maternal genetic effects were about one-third as large as direct genetic effects. A prenatal line × postnatal line interaction for weight and percentage of gonadal fat was caused by the inability of S line dams to provide sufficient milk to maintain normal development of L line young. Further evidence supporting this hypothesis was the high mortality rate among L line young reared by S line dams when compared with the mortality in all other subgroups. Increasing postnatal litter size reduced weight and percentage of gonadal fat, but this factor was of less importance than direct genetic effects. In general, the relative importance of direct genetic, postnatal maternal genetic and postnatal litter size effects was similar for metric measures of growth (body weight, body length and tail length) and for adiposity (gonadal fat pad weight and percentage). Rate of gonadal fat pad development relative to body weight was higher in line L than in line S. At a constant body weight, however, line L mice had less fat than line S mice.

(Key Words: Maternal Effects, Fat Deposition, Correlated Responses, Mice.)

Introduction

Selection for large or small body weight in the mouse results in correlated responses in maternal performance that influence growth of the young (Legates, 1972; Eisen, 1974). Maternal effects occur prenatally and postnatally, and include both genetic and environmental factors. Prenatal maternal effects are dependent on the uterine environment, while postnatal maternal effects are determined by lactational output and behavior of the mother. Selection for body weight generally results in a positive correlated response in litter size, which causes an indirect negative maternal effect on prenatal and preweaning growth of the progeny. To separate correlated responses in maternal effects on growth and fat deposition from direct effects of genes possessed by the young, researchers have introduced experimental designs that utilize egg transfer, crossfostering and standardization of litter size (Brumby, 1960; White *et al.*, 1968; Al-Murrani and Roberts, 1978; Hayes and Eisen, 1979a).

Egg transfers between lines of mice selected for large and small body size have shown that the prenatal maternal environment is superior in large mothers, but that the effect on body weight is relatively small at birth and is almost completely absent by 2 weeks of age (Moore *et al.*, 1970; Al-Murrani and Roberts, 1978). Crossfostering studies with large and small lines of mice have demonstrated that young nursed by large mothers have higher preweaning body weights than young nursed by small mothers; after weaning, the difference in body weight is maintained (White *et al.*, 1968).

[1] Paper No. 6452 of the Journal Ser. of the North Carolina Agr. Res. Service, Raleigh. Supported, in part, by NIH Research Grant AM21950.

[2] Permanent address: Dept. of Anim. Sci., North Carolina State Univ., Raleigh 27650.

[3] Supported, in part, by an Underwood Fund Fellowship, Agricultural Research Council, and an Edinburgh University Biological Fellowship in the Dept. of Genetics.

[4] Agricultural Research Council, Unit of Animal Genetics.

When large body weight lines have been compared with unselected controls, the correlated response in postnatal maternal performance on preweaning growth has varied from positive (La Salle and White, 1975; Nagai, 1977; Hayes and Eisen, 1979b) to negligible values (White *et al.*, 1968; Nagai *et al.*, 1976). While litter size at birth has a negative effect on birth weight (Al-Murrani and Roberts, 1978; Eisen and Durrant, 1980), its effect on postnatal maternal performance in large and small lines is negligible when progeny are reared in standardized litter sizes (Eisen *et al.*, 1980). In contrast, the effect on body weight of varying postnatal litter size persists throughout the period of postweaning growth in lines selected for high or low body weight (Eisen and Leatherwood, 1978; Hayes and Eisen, 1979a).

Selection for divergence in growth rate or body weight generally results in positive correlated responses in percentage body fat (Eisen, 1974). However, there is a paucity of data on the influence of maternal effects on fat deposition. Hayes and Eisen (1979b) reported a large positive correlated response in postnatal maternal effects on weight and percentage body fat at the peak of lactation (12 days of age) when comparing lines selected for high and low 6-week body weight. Whether the postnatal maternal genetic effects on body fat are still present at postweaning ages has not been determined. In lines selected for high or low body weight, postweaning body weight and fat weight decreased permanently as postnatal litter size increased (Eisen and Leatherwood, 1978; Hayes and Eisen, 1979b). Because of interactions, if differences among selected lines in percentage body fat are small, the effects of postnatal litter size are difficult to predict (Hayes and Eisen, 1979b). In contrast, selected lines having a large correlated response in percentage fat continue to show the effects of postnatal litter size on percentage body fat to at least 30 weeks of age (Eisen and Leatherwood, 1978).

The objectives of the present study were to assess the relative importance of the following factors on growth and fat deposition in lines of mice selected for large and small body size: (1) correlated responses in direct genetic effects, (2) correlated responses in postnatal maternal performance, (3) postnatal litter size effects and (4) interactions among these factors.

Materials and Methods

Mice used were samples of the replicate Q lines, selected by Falconer (1973) for high and low 6-week body weight. Two large and two small lines were available. Selection had been relaxed for 21 to 24 generations before the present experiment. Lines of the same size differed by a color marker gene (albino *versus* colored).

There were two replicates of crossfostered litters. The large colored and small albino lines formed one replicate, and the large albino and small colored lines composed the second. Only first litters were used. Figure 1 helps one visualize the experimental design. On the day of birth (day 0), progeny of two dams from each of the large (L) and small (S) lines formed a "set," the four dams having given birth to a minimum of six offspring each within the same 24-hr period. The two dams of each line were randomly assigned to rear either four or eight young. Allocation of mice to each dam was done at random, with an attempt made to keep a balance between male and female offspring. A treatment subclass was designated by prenatal line (L_g or S_g),

CROSSFOSTERED SET

		L_m		S_m	
		N_4	N_8	N_4	N_8
L_g	D_1	1	2	1	2
	D_2	1	2	1	2
S_g	D_3	1	2	1	2
	D_4	1	2	1	2

L_g, S_g = PRENATAL LINES

$D_1, \ldots, D_4$ = PRENATAL DAMS

L_m, S_m = POSTNATAL LINES

N_4, N_8 = POSTNATAL LITTER SIZES

Figure 1. Diagrammatic representation of the crossfostering experiment.

postnatal line (L_m or S_m) and postnatal litter size (4 or 8); e.g., L_gS_m4 refers to large mice reared by small dams in litters of four.

Mice were toe-clipped on day 0 for identification of the prenatal dam within each line, and again at 12 days of age for individual identification. Offspring were weaned at 3 weeks and caged by prenatal line × postnatal line × postnatal litter size × sex subgroups. Individual body weights were recorded at birth, 12 days and 3 and 6 weeks of age.

Absolute growth rate, $(W_i-W_j)/t$, and relative growth rate, $100(\ln W_i-\ln W_j)/t$, were calculated for adjacent weight periods, where W_i = final weight, W_j = initial weight and t = number of days between recording of weights. At 6 weeks of age, naso-anal body length and tail length were recorded. The Lee index, $10(W_6)^{1/3}/L_6$, was calculated, where W_6 = 6-week body weight and L_6 = 6-week body length. The animals were then killed, and both gonadal fat pads were excised and weighed. Gonadal fat was also expressed as a percentage of body weight.

The gonadal fat pads were used as an indirect measure of total fat, as facilities for the determination of total fat were not available. The correlation between gonadal fat and total body fat is sufficiently high to justify the use of this procedure (Eisen and Leatherwood, 1978; Jagot *et al.*, 1980; Rogers and Webb, 1980). The Lee index was used as an index of obesity (Bernardis, 1970; Dubuc, 1976).

The data were analyzed by least-squares procedures for unequal subclass numbers (Harvey, 1975). Prenatal line effects were tested with the prenatal dam mean square in the analysis of variance. The other factors of interest, which included postnatal line, postnatal litter size, sex and respective interaction effects, were tested with the residual mean square.

The prenatal line effect includes direct genetic effects as well as prenatal maternal genetic effects and the direct × prenatal maternal interaction. Since Al-Murrani and Roberts (1978) found that the last two effects were not important in the L and S lines, they were assumed to be zero.

A covariance analysis was conducted to compare treatment differences in gonadal fat pad weight at a constant body weight. Gonadal fat pad weight and body weight were transformed to logarithms to ensure a linear allometric relationship, as recommended by Hayes and McCarthy (1976).

Results

Replicates 1 and 2 provided 21 and 15 crossfostered sets. While replicate and sex differences were significant for many traits, interactions involving these factors were generally small. Therefore, means presented are pooled over replicates and sexes.

Characteristics of the Postnatal Dam. Means for traits of the postnatal dam are given in table 1. The results provide a check on the success of the randomization used in the formation of the crossfostered sets. Dams of the same line assigned to rear either four or eight young did not differ significantly in litter size or body weight on day 0. Dams of the L and S lines assigned to rear the same number of young had identical crossfostered litter weights on day 0. In addition, the correlations between crossfostered litter weight on day 0 and litter size or dam body weight on day 0 were not significantly different from zero (table 2). Our conclusion is that the goal of the randomization procedure used on day 0 was realized.

Dams of the L line had a larger ($P<.01$) litter size at birth than did dams of the S line, although both means were biased upward slightly because dams with litters of fewer than six young were not considered for crossfostering. Dams nursing eight young had a higher ($P<.01$) relative growth rate during lactation than did those suckling four, while dams of the S line had a higher ($P<.01$) relative growth rate than did those of the L line. Eisen *et al.* (1977) found that mammary gland development increased with increasing postnatal litter size, which may account for the increased size of the dams suckling a greater number of pups. Eisen and Durrant (1980) also found that relative growth rate of dams during lactation was greater in a line selected for small body weight than in a large line.

Relative growth rate of the litter from 0 to 12 days was higher both for litters from large mothers and for small litters, each effect being of similar magnitude. This suggests that if postnatal litter size had not been standardized, the better postnatal maternal genotype of line L would have been offset almost entirely by the positive correlated response in litter size

TABLE 1. MEANS AND LINEAR CONTRASTS FOR TRAITS OF THE POSTNATAL DAM

Treatment	No. of dams	Litter size at birth	Dam body weight, g		Dam relative growth rate, %	Crossfostered litter weight, g		Litter relative growth rate, %
			Day 0	Day 12		Day 0	Day 12[e]	
L_m4	36	11.2^a	45.9^a	47.7^a	$.35^a$	6.3^a	37.5^a	14.8^a
L_m8	36	11.5^a	47.0^a	52.1^b	$.86^b$	12.6^b	59.8^b	12.9^b
S_m4	36	8.6^b	24.9^b	27.4^c	$.79^b$	6.3^a	28.5^c	12.6^b
S_m8	36	8.6^b	24.5^b	29.0^c	1.40^c	12.6^b	42.8^d	10.1^c
SE		.37	.64	.63	.09	.10	.82	.17
Contrast ± SE								
Postnatal line (M) (L_m-S_m)		2.7 ± .4**	21.7 ± .6**	21.7 ± .6**	−.49 ± .09**	.0 ± .1	13.0 ± .8**	2.5 ± .2**
Postnatal litter size (N) (4 − 8)		−.2 ± .4	−.4 ± .6	−3.0 ± .6**	−.56 ± .09**	−6.3 ± .1**	−18.3 ± .8**	2.1 ± .2**
M × N		−.3 ± .7	−1.5 ± 1.3	−2.7 ± 1.2*	.10 ± .18	.0 ± .2	−8.0 ± 1.6**	−.6 ± .4

[a,b,c,d]Means in the same column under the same heading with no common superscripts differ (P<.05).

[e]All litter weights were adjusted to a litter size of four or eight pups to compensate for mortality occurring between days 0 and 12.

*P<.05.

**P<.01.

TABLE 2. CORRELATIONS AMONG TRAITS OF THE POSTNATAL DAM[a]

Trait		(2)	(3)	(4)	(5)
Litter size at birth	(1)	−.10	.12	.40**	.33**
Crossfostered litter weight; day 0	(2)		.21*	−.02	.05
Crossfostered litter weight; day 12	(3)			.24**	.24**
Dam body weight; day 0	(4)				.83**
Dam body weight; day 12	(5)				

[a]Pooled within postnatal line by postnatal litter size subclasses (df = 136).
*P<.05.
**P<.01.

of line L, and *vice versa* for line S. This prediction, however, depends on the assumptions of (1) linearity of the regression of relative litter growth on postnatal litter size, (2) absence of a line × postnatal litter size interaction, (3) symmetrical correlated responses in maternal performance and litter size and (4) no differential line mortality when females rear their natural litters. These points will apply equally to the individual measurements of progeny growth in the sections that follow.

The correlation between 12-day crossfostered litter weight and litter size at birth was not significantly different from zero, while the positive correlation between 12-day crossfostered litter weight and dam body weight on day 0 was significant (table 2). This agrees with the findings of Eisen *et al.* (1980) that increased fetal mass does not enhance lactational output by the dam. However, Skjervold (1977) did report a positive association between fetal mass and postnatal maternal performance.

Mortality Data. Large mice nursed by small dams suffered a high preweaning mortality rate, 20% (table 3). Postweaning mortality rate among these large line mice abated when they were reared in litters of four, but continued at the same high rate when they were reared in litters of eight. Small young reared by the same line S dams as the large young showed a cumulative mortality rate of only 6%, ruling out the possibility of disease or behavioral defects in these mothers. We conclude that females of line S were not able to provide sufficient milk for line L young, the effect being more severe as litter size increased.

Preweaning Growth. Means and linear contrasts of preweaning growth traits are presented in table 4. Prenatal (line of origin) effects were significant (P<.01) for body weights from birth to weaning and preweaning weight gains, the L line being heavier, as expected. This was a consequence of the positive correlated responses in preweaning growth resulting from selection for 6-week body weight. As expected, only the pre-

TABLE 3. MORTALITY RATES AMONG PROGENY

	Preweaning		Postweaning		Total
Treatment	No. at day 0	% dead	No. at day 21	% dead	% dead
L_gL_m4	72	6.9 ± 3.0	67	1.5 ± 1.5	8.3 ± 3.3
L_gL_m8	144	4.9 ± 1.8	137	5.8 ± 2.0	10.4 ± 2.5
L_gS_m4	72	19.4 ± 4.7	58	6.9 ± 3.0	25.0 ± 5.1
L_gS_m8	144	20.8 ± 3.4	114	18.4 ± 3.9	35.4 ± 4.0
S_gL_m4	72	2.8 ± 1.9	70	2.9 ± 2.0	5.6 ± 2.8
S_gL_m8	144	10.4 ± 2.5	129	.0 ± .0	10.4 ± 2.6
S_gS_m4	72	2.8 ± 1.9	70	2.9 ± 2.0	5.6 ± 2.8
S_gS_m8	144	5.6 ± 1.9	136	.7 ± .7	6.3 ± 1.9

TABLE 4. MEANS ± SE AND LINEAR CONTRASTS OF GROWTH TRAITS FROM BIRTH TO WEANING

Treatment	Body weight, g			Gain, g/day		Relative growth rate, %	
	Birth	12 days	3 weeks	Birth to 12 days	12 days to 3 weeks	Birth to 12 days	12 days to 3 weeks
L_gL_m4	1.78 ± .01	10.58 ± .10	16.76 ± .19	.73 ± .008	.69 ± .014	14.82 ± .13	5.18 ± .14
L_gL_m8	1.77 ± .02	8.35 ± .07	13.51 ± .13	.55 ± .006	.57 ± .010	12.89 ± .09	5.23 ± .10
L_gS_m4	1.76 ± .01	7.44 ± .11	10.93 ± .20	.47 ± .008	.38 ± .015	11.95 ± .13	3.96 ± .15
L_gS_m8	1.76 ± .02	5.74 ± .07	7.95 ± .14	.33 ± .006	.24 ± .011	9.75 ± .09	3.31 ± .11
S_gL_m4	1.38 ± .01	8.33 ± .10	11.48 ± .18	.58 ± .008	.35 ± .014	14.89 ± .13	3.58 ± .14
S_gL_m8	1.39 ± .02	6.57 ± .08	9.21 ± .14	.43 ± .006	.29 ± .010	12.80 ± .10	3.62 ± .10
S_gS_m4	1.37 ± .01	6.73 ± .10	9.45 ± .18	.45 ± .008	.30 ± .014	13.21 ± .13	3.74 ± .14
S_gS_m8	1.38 ± .02	4.93 ± .07	6.85 ± .13	.30 ± .006	.21 ± .010	10.54 ± .09	3.51 ± .10

Contrast	Diff.	%[a]	Diff.	%	Diff.	%	Diff.	%	Diff.	%	Diff.	%	Diff.	%
Prenatal line (G) (L_g-S_g)	.39** (.02)[b]	24.7	1.39** (.06)	19.0	3.04** (.16)	28.2	.08** (.005)	16.7	.18** (.010)	47.5	−.51** (.09)	−4.0	.81** (.08)	20.2
Postnatal line (M) (L_m-S_m)	.01	.6	2.25**	30.7	3.95**	36.6	.19**	39.7	.19**	50.1	2.49**	19.8	.77**	19.2
Postnatal litter size (N) (4 − 8)	.00	.0	1.87**	25.5	2.77**	25.7	.16**	33.4	.11**	29.0	2.23**	17.7	.20*	5.0
G X M	.00	.0	.63**	8.6	1.75**	16.2	.05**	10.4	.13**	34.3	.52**	4.1	.80**	19.9
G X N	.01	.6	.09	1.2	.34**	3.2	.007	1.5	.027**	7.1	−.16*	1.3	.10	2.5
M X N	.00	.0	.11	1.5	−.01	−.1	.009	1.9	−.010	−2.6	−.21**	1.7	−.24**	−6.0
G X M X N	.00	.0	.14*	1.9	.15	1.4	.012*	2.5	.002	.1	.08	0.6	−.11	−2.7
Sex (♂♂−♀♀)	.01	.6	.03	.4	.25*	2.3	.000	.0	.029**	7.7	−.16*	−1.3	.25**	6.2
SE of contrast[c]	.01		.06		.12		.005		.009		.08		.08	

[a]Difference as a percentage of the overall mean.

[b]Values in parentheses are standard errors of prenatal line contrast.

[c]Standard error of all contrasts except the prenatal line contrast.

*P<.05.

**P<.01.

natal line effect was significant for birth weight; the difference favoring the L line was 25% of the overall mean. Correlated responses in postnatal line effects on preweaning weights and gains also favored line L. Postnatal effects exceeded prenatal effects by 62% for 12-day weight and by 138% for gain from birth to 12 days. Postnatal effects were only 30% greater than prenatal effects for 3-week body weight. The four traits, 12-day and 3-week body weight and gain over the two periods, exhibited a significant ($P<.01$) prenatal line × postnatal line interaction. The basis for the interaction is illustrated in figure 2, which shows the subclass means for various traits. Differences between L and S offspring in weights and particularly in weight gains were much greater when the mice were nursed by L line mothers rather than by S line mothers.

Postnatal litter size effects were equivalent to declines in 12-day weight and 3-week body weight of .47 and .69 g for each additional pup being nursed. The linear contrast ($L_m - S_m$) in table 4 shows a postnatal superiority of 3.95 g for the L line in individual 3-week weight of the offspring when litter sizes were equal. But since L line mothers suckling natural litters would have had three more pups to rear (table 1), these pups would have had their weights reduced by 3 × .69 g, or by about one-half of the potential postnatal superiority of large mothers. The postnatal line × postnatal litter size interactions for 3-week weight and 12-day to 3-week gain were small (figure 3), and although statistically significant, they appear to be biologically unimportant. The significant three-factor interactions for 12-day weight and birth to 12-day gain were also unimportant.

Relative growth rate describes the logarithmic change in body weight per day. For relative growth rate, the prenatal effects of the S line exceeded the prenatal effects of the L line from birth to 12 days, but the ranking was reversed from 12 days to 3 weeks. Postnatal line effects for relative growth rate favored line L. However, the significant prenatal line × postnatal line interactions influenced the ranking of lines (figure 2) in relative preweaning growth rate. When L and S line offspring were nursed by L line dams, they had similar relative growth rates from birth to 12 days, whereas when they were reared by S line dams, the S line offspring had a greater relative growth rate. But from 12 days to 3 weeks, the relative growth rate of S line offspring was the same whether they were reared by L or S dams, whereas L line offspring had a reduced relative growth rate when reared by S rather than L line dams. The results reinforce the conclusion, stated earlier, that S line dams do not provide sufficient milk to sustain the growth needs of the L young.

Progeny reared in litters of four had a higher ($P<.01$) preweaning relative growth rate than did those reared in litters of eight. Prenatal line × postnatal litter size and postnatal line × postnatal litter size interactions for preweaning relative growth rate, although statistically significant, did not result in a change in the ranking of lines with different litter sizes (figures 3 and 4).

Postweaning Growth. The prenatal line means for the five characteristics of postweaning growth were much greater for the L than for the S line (table 5), reflecting the genetic effects of line of origin. The differences, as proportions of the overall means, were 56, 15 and 11% for 6-week body weight, body length and tail length, respectively. Body weight of individuals of the L genotype exceeded that of individuals of the S genotype by about 79% at 6 weeks, which agrees with the value reported by Al-Murrani and Roberts (1978) for the same genetic stocks. Postnatal line effects favoring the large line mothers were carried over to 6 weeks of age, but their importance in relation to prenatal line effects was only 25% for body weight, 30% for body length and 40% for tail length. Compensatory postweaning growth was observed among progeny reared by line S mothers; they had a higher 3- to 6-week relative growth rate than did progeny reared by L line mothers. The prenatal line × postnatal line interactions for the postweaning growth traits had no effect on the ranking of genotypes (figure 2), and, hence, they can safely be ignored. Although significant, the interactions were probably only a carry-over effect of the preweaning influence of the S line dams nursing L line young.

Body weight, body length and tail length were greater for progeny reared in litters of four than for those reared in litters of eight. The absence of a difference in postweaning gain indicates that the postnatal litter size effect on postweaning growth occurred entirely before weaning. Compensatory postweaning

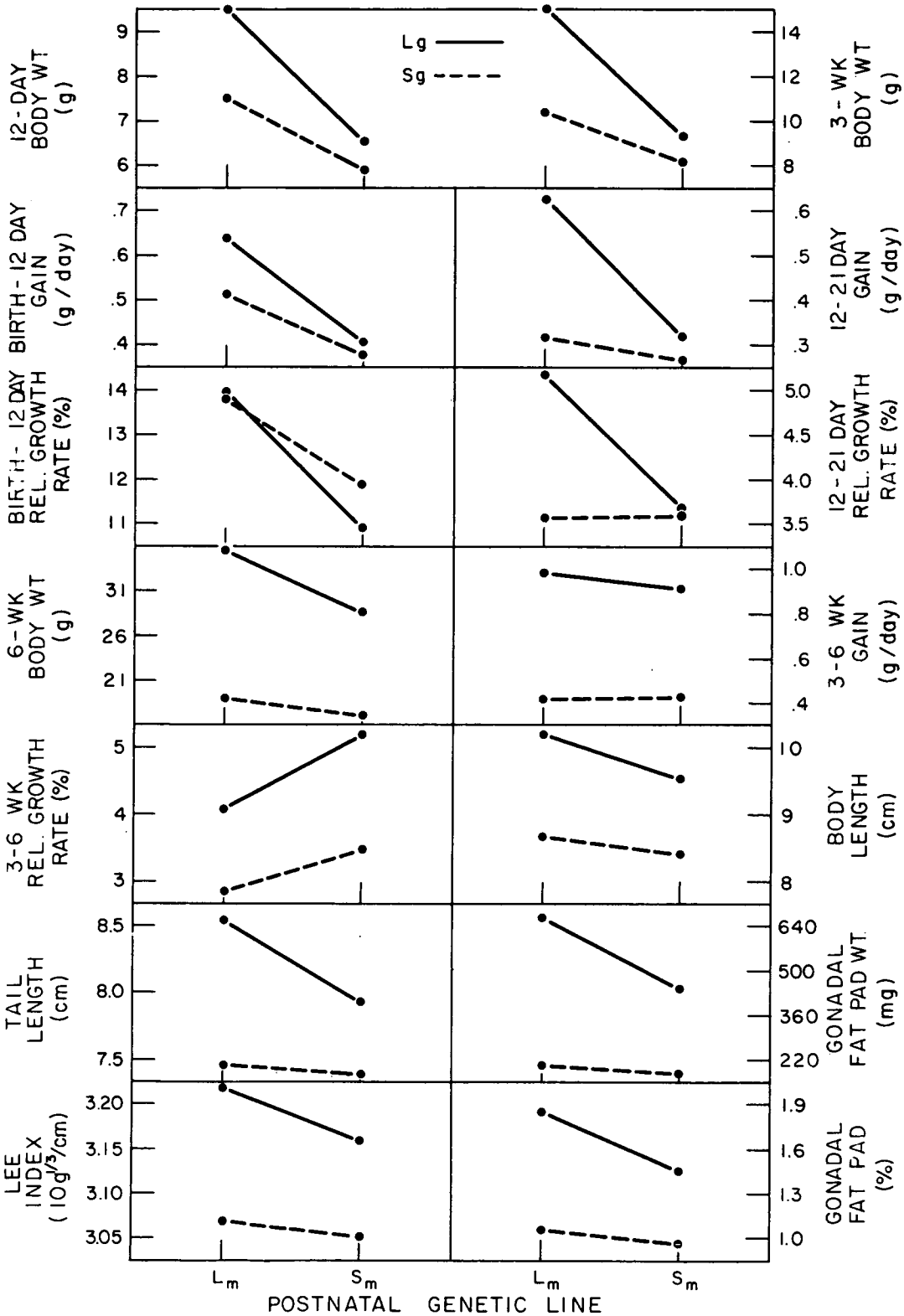


Figure 2. Plot of prenatal line X postnatal line subclass means for traits having a statistically significant interaction.

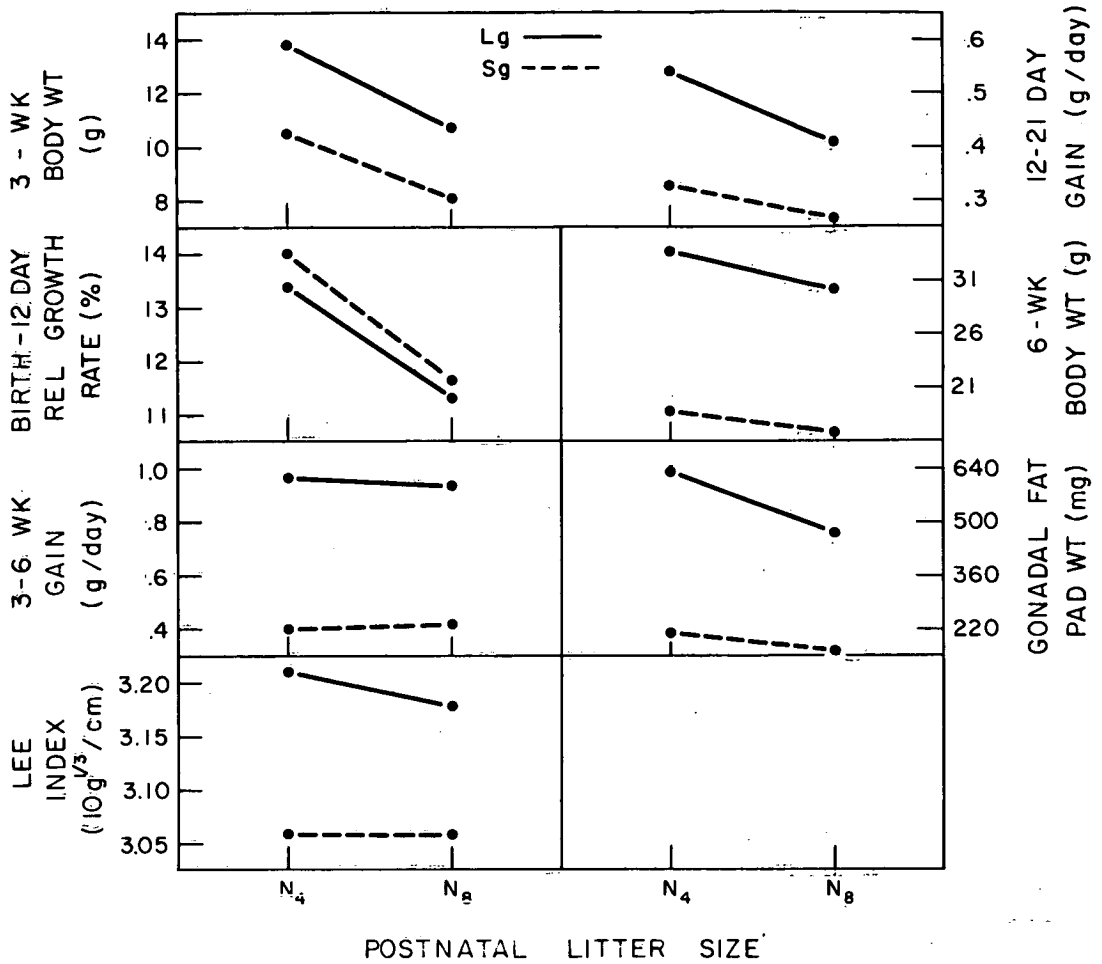


Figure 3. Plot of prenatal line X postnatal litter size subclass means for traits having a statistically significant interaction.

relative growth rate was detected among progeny reared in litters of eight. The significant prenatal line X postnatal litter size and postnatal line X postnatal litter size interactions were of only minor importance (figures 3 and 4).

Fat Deposition. Direct genetic effects of line L exceeded those of line S for the Lee index and weight and percentage of gonadal fat at 6 weeks of age (table 6). The postnatal line effects for the indices of body fat showed that the postnatal maternal impact was important, with the maternal environment of the large line resulting in a fatter mouse 3 weeks after weaning. Postnatal line effects were about one-third as large as prenatal line effects for measures of fat deposition at 6 weeks.

The prenatal line X postnatal line interaction was significant for the Lee index and for both weight and percentage of gonadal fat at 6 weeks. Figure 2 illustrates that the interaction was caused by a greater decline in the mean for large mice than in that for small mice when both lines were nursed by line S dams.

Postnatal litter size effects also were significant for the indices of fat deposition at 6 weeks. Fat pad weight and percentage declined by 26 mg and .065% with each one-pup increase in postnatal litter size.

The significant prenatal line X postnatal litter size interaction for the Lee index and gonadal fat pad weight did not affect line rankings (figure 3). None of the indices of fat content was significantly affected by a postnatal line X postnatal litter size interaction.

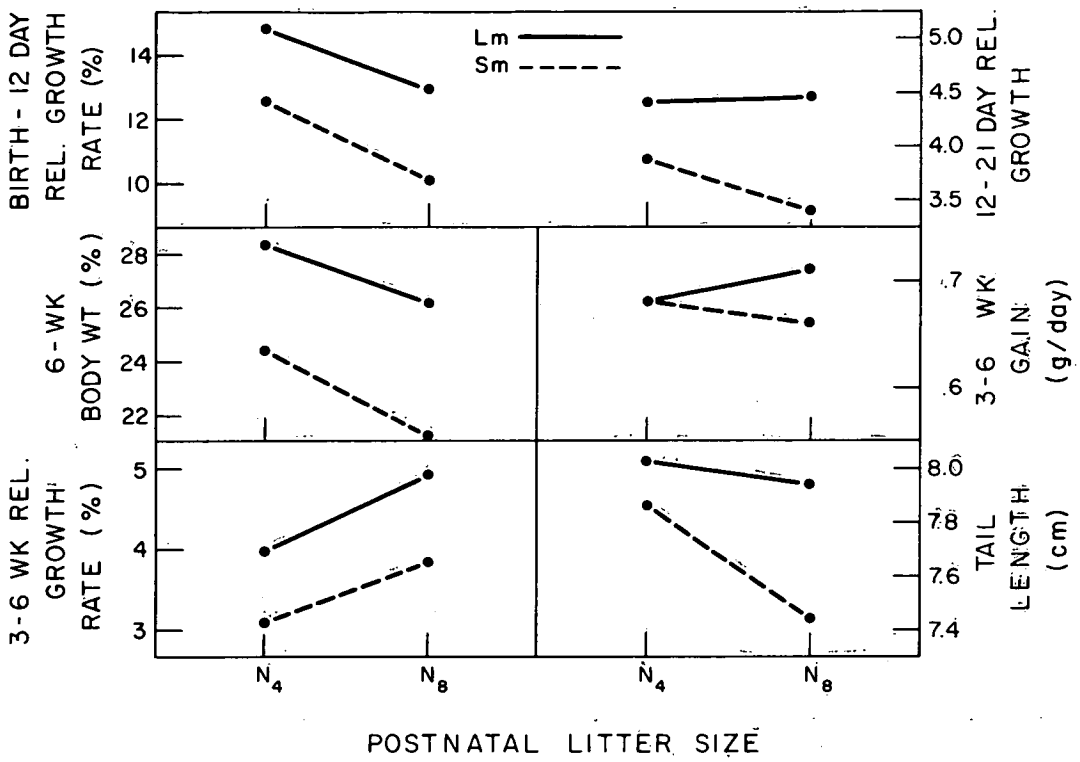


Figure 4. Plot of postnatal line X postnatal litter size subclass means for traits having a statistically significant interaction.

The regression lines of log gonadal fat pad weight on log body weight, shown in figure 5, were heterogeneous in slope ($F_{7,674} = 5.58$, $P<.01$) and therefore cannot be tested formally for differences in elevation. Nevertheless, the slopes for the four S_g groups clearly had a higher elevation than those for the four L_g groups over the range of the data. Whenever the body weights overlapped, mice in the S_g group unambiguously had more fat than those in any of the L_g groups. Neither the line of the foster dam nor the postnatal litter size appeared to contribute to differences in fat content at a constant body weight.

The heterogeneous slopes of log gonadal fat pad weight on log body weight were compared on the basis of the same linear contrasts used for assessing mean differences. The slope for the L line was greater ($P<.05$) than that for the S line for prenatal effects ($3.21 \pm .16$ *vs* $2.76 \pm .14$). The rate of development of gonadal fat relative to body weight at 6 weeks of age thus appears to have been higher in the L line than in the S line.

A significant ($P<.01$) postnatal line x postnatal litter size interaction for the slopes made it difficult to interpret the importance of these two effects. The regression coefficients were L_m4, $3.26 \pm .28$; L_m8, $2.24 \pm .19$; S_m4, $2.88 \pm .18$; and S_m8, $3.56 \pm .12$. Thus, the rate of fat development relative to body growth was greater with line L mothers than with line S mothers in litters of four, but the ranking of the foster mothers was reversed in litters of eight.

Correlations among selected traits measured at 6 weeks are given in table 7. The Lee index had a low correlation with gonadal fat pad weight and percentage. The results agree with earlier findings which showed that the Lee index is not a good predictor of either total body fat (Bernardis, 1970) or gonadal fat as a percentage of body weight (Rogers and Webb, 1980). Fat pad weight and percentage were highly correlated. Body weight and body length were highly correlated, but body length was slightly less correlated with fat pad weight or percentage than was body weight. Tail length had a low correlation with fat pad weight and percentage.

TABLE 5. MEANS ± SE AND LINEAR CONTRASTS OF POSTWEANING GROWTH TRAITS

Treatment	6-week body weight, g	3- to 6-week gain, g/day	3- to 6-week relative growth rate, %	6-week body length, cm	6-week tail length, cm
L_gL_m4	37.17 ± .35	.97 ± .014	3.78 ± .07	10.32 ± .04	8.57 ± .09
L_gL_m8	34.36 ± .25	.99 ± .010	4.47 ± .05	10.10 ± .03	8.46 ± .07
L_gS_m4	30.72 ± .39	.94 ± .015	4.99 ± .08	9.79 ± .05	8.17 ± .11
L_gS_m8	26.67 ± .30	.87 ± .012	5.76 ± .06	9.46 ± .04	7.67 ± .08
S_gL_m4	19.58 ± .35	.39 ± .013	2.51 ± .07	8.79 ± .04	7.48 ± .10
S_gL_m8	18.18 ± .26	.43 ± .010	3.25 ± .05	8.54 ± .03	7.39 ± .07
S_gS_m4	18.06 ± .35	.41 ± .014	3.05 ± .07	8.58 ± .04	7.54 ± .10
S_gS_m8	16.08 ± .25	.44 ± .010	4.10 ± .05	8.27 ± .03	7.19 ± .07

Contrast	Diff.	%[a]	Diff.	%	Diff.	%	Diff.	%	Diff.	%
Prenatal line (G) (L_g-S_g)	14.26** (.26)[b]	56.8	.53** (.011)	77.9	1.52** (.06)	38.1	1.37** (.03)	14.8	.82** (.06)	10.5
Postnatal line (M) (L_m-S_m)	3.57**	14.2	.03**	4.4	−.97**	−24.3	.41**	4.4	.33**	4.2
Postnatal litter size (N) (4 − 8)	2.56**	10.2	−.01	−1.5	−.81**	−20.3	.28**	3.0	.26**	3.3
G X M	2.63**	10.5	.04**	5.9	−.28**	−7.0	.17**	1.8	.26**	3.3
G X N	.87**	3.5	.03**	4.4	.08	2.0	.00	.0	.05	.6
M X N	−.46*	−1.8	−.02*	−2.9	.10*	2.5	−.05	−.5	−.16*	−2.1
G X M X N	−.16	−.6	−.01	−1.5	−.05	−1.3	−.01	−.1	−.03	−.4
Sex (♂♂−♀♀)	4.03**	16.1	.18**	26.5	.64**	16.1	.27**	2.9	.01	.1
SE of contrast[c]	.22		.009		.05		.03		.06	

[a]Difference as a percentage of the overall mean.

[b]Values in parentheses are standard errors of prenatal line contrasts.

[c]Standard error of all contrasts except that for the prenatal line contrast.

*P<.05.

**P<.01.

Discussion

The results clearly demonstrate that selection for 6-week body weight in mice results in positive correlated responses in postnatal maternal genetic effects, in addition to direct genetic effects, on growth rate and fat deposition. The relative importance of postnatal maternal genetic effects and direct genetic effects on fat deposition varies with age. Correlated responses in postnatal maternal genetic and direct genetic effects on fat content follow the same pattern that has been observed for growth traits in this and other studies (White *et al.*, 1968; La Salle and White, 1975; Nagai *et al.*, 1976; Nagai, 1977). Hayes and Eisen (1979b) showed that postnatal maternal genetic effects are more important than direct genetic effects in explaining the greater weight and percentage of body fat at 12 days of age in large line than in small line young. As the present study indicates, after weaning, the importance of direct genetic effects overtakes and surpasses that of postnatal maternal genetic effects. Thus, at 6 weeks of age, the direct genetic effects account for considerably more of the difference in weight and percentage of the gonadal fat pad than do postnatal maternal genetic effects. Nevertheless, postnatal maternal genetic effects are still sizable 3 weeks after weaning, being about one-third as large as direct genetic effects.

Although line L exceeded line S in direct genetic and postnatal maternal genetic effects on all body weights and postweaning fat deposition, a clear prenatal line x postnatal line interaction was apparent. Hayes and Eisen (1979b) found such an interaction for fat deposition prior to weaning. The interaction can be described by the greater superiority of line L dams over line S dams when each reared line L young than when each reared line S young; i.e., $(L_gL_m - L_gS_m) > (S_gL_m - S_gS_m)$. The results support the earlier findings

TABLE 6. MEANS ± SE AND LINEAR CONTRASTS OF MEASURES OF FAT DEPOSITION AT 6 WEEKS OF AGE

Treatment	Lee index, 10 $g^{1/3}$/cm	Gonadal fat pad weight, mg	Gonadal fat pad weight/body weight, %
L_gL_m4	3.23 ± .008	740 ± 22	1.99 ± .06
L_gL_m8	3.21 ± .006	589 ± 16	1.70 ± .04
L_gS_m4	3.18 ± .009	527 ± 24	1.62 ± .07
L_gS_m8	3.14 ± .007	368 ± 18	1.30 ± .05
S_gL_m4	3.06 ± .008	229 ± 21	1.15 ± .06
S_gL_m8	3.07 ± .006	187 ± 16	.99 ± .05
S_gS_m4	3.05 ± .008	199 ± 22	1.07 ± .06
S_gS_m8	3.04 ± .006	131 ± 16	.78 ± .04

Contrast	Diff.	%	Diff.	%	Diff.	%
Prenatal line (G) (L_g-S_g)	.14** (.006)[b]	4.5	369** (17)	99.4	.65** (.05)	49.1
Postnatal line (M) (L_m-S_m)	.04**	1.3	130**	35.0	.26**	19.6
Postnatal litter size (N) (4 − 8)	.01*	.3	105**	28.3	.26**	19.6
G X M	.02**	.6	87**	23.4	.12**	9.1
G X N	.02**	.6	50**	13.5	.04	3.0
M X N	.00	.0	−8	−2.2	−.04	−3.0
G X M X N	.00	.0	4	1.1	.03	2.3
Sex (♂♂−♀♀)	.07**	2.2	−21**	5.7	−.19**	−14.3
SE of contrast[c]	.005		14		.04	

[a]Difference as a percentage of the overall mean.

[b]Values in parentheses are standard errors of prenatal line contrasts.

[c]Standard error of all contrasts except that for the prenatal line contrast.

*P<.05.

**P<.01.

of inadequate postnatal maternal performance by line S dams when nursing line L offspring (Al-Murrani and Roberts, 1978). The high mortality rate among line L progeny reared by line S mothers is further evidence that line S mothers are unable to provide an adequate supply of milk to line L young. The severity of the interaction effect probably was not as great in the lines studied by Hayes and Eisen (1979b), because no differential mortality was apparent up to 12 days.

Increasing postnatal litter size reduced preweaning growth, as expected, and also reduced postweaning body weight and gonadal fat pad weight and percentage. Postnatal litter size effects on postweaning growth and fat deposition were of less importance than direct genetic effects. Eisen and Leatherwood (1978) reported a decrease in adiposity as postnatal litter size was increased from eight to 14 pups in a line selected for high postweaning gain. Hayes and Eisen (1979a) also reported that fat deposition tended to decline as postnatal litter size increased in lines selected for high and low body weight.

Selection for 6-week weight led to a positive correlated response in direct genetic effects on the rate of development of gonadal fat relative to body weight. Eisen and Leatherwood (1978) also noted that selection for rapid growth yielded a positive correlated response in the regression of log fat pad weight on log body weight. In contrasting the regressions of log body fat weight on log body weight, Hayes and McCarthy (1976) found that when selection for body weight was conducted at 10 weeks, the high line exceeded the low line, whereas, when selection was at 5 weeks, the lines did not differ.

When comparisons were made at a constant body weight, the direct genetic effect for gonadal fat pad weight for the large line was less than that for the small line. Hayes and McCarthy (1976) did not find a correlated

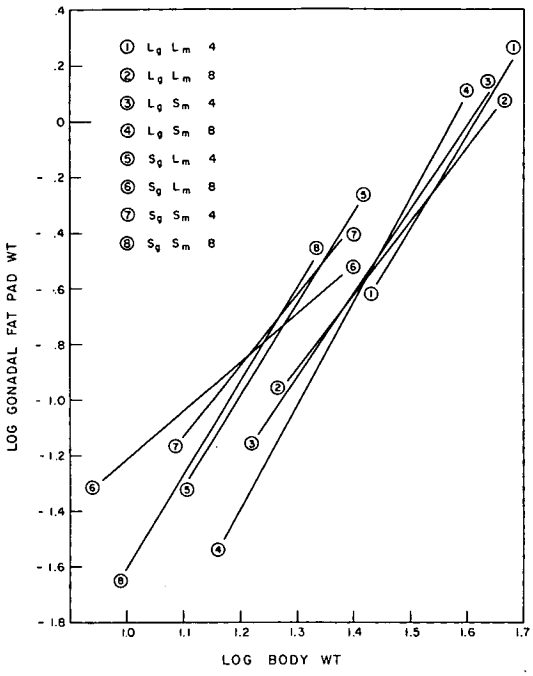


Figure 5. Regression lines of log gonadal fat pad weight on log body weight at 6 weeks of age.

response in fat weight at a constant body weight when selection was for high or low 5-week weight; the slopes were heterogeneous in the lines selected for weight at 10 weeks so that mean contrasts were not conducted.

Eisen and Leatherwood (1978) found that a line selected for high body weight exceeded an unselected control in total body fat and in cell size and cell number of the gonadal fat pad. Several studies have shown positive correlated responses in fat percentage as a result of selection for postweaning body weight or rate of gain (e.g., Hayes and McCarthy, 1976). Only the present study and that of Hayes and Eisen (1979b) have partitioned the correlated response in fat deposition into direct genetic and postnatal maternal genetic components.

In conclusion, positive correlated responses in postnatal maternal genetic effects, generally associated with lactational performance of the mother, can contribute significantly to differences in degree of adiposity in mice selected for rapid growth rate. The positive correlated response in fat deposition due to postnatal maternal effects may be offset, in part, if number born has increased as a correlated response and litters are not standardized at birth. Whether these effects are important in livestock selected for rapid postweaning growth is presently a moot question.

TABLE 7. CORRELATIONS AMONG RESIDUALS FOR TRAITS MEASURED AT 6 WEEKS[a]

Trait		(2)	(3)	(4)	(5)	(6)
Body weight	(1)	.82	.59	.48	.37	.33
Body length	(2)		.43	.37	−.16	.31
Gonadal fat pad weight	(3)			.95	.26	.21
Gonadal fat weight/ body weight	(4)				.22	.20
Lee index	(5)					.10
Tail length	(6)					

[a]Degrees of freedom = 683. All correlations are significant at $P<.01$, except that between Lee index and tail length ($P<.05$).

Literature Cited

Al-Murrani, W. K. and R. C. Roberts. 1978. Maternal effects on body weight in mice selected for large and small size. Genet. Res. 32:295.

Bernardis, L. L. 1970. Prediction of carcass fat, water and lean body mass from Lee's 'nutritive ratio' in rats with hypothalamic obesity. Experientia 26:789.

Brumby, P. J. 1960. The influence of the maternal environment on growth in mice. Heredity 14:1.

Dubuc, P. U. 1976. Effects of limited food intake on the obese-hyperglycemic syndrome. Amer. J. Physiol. 230:1474.

Eisen, E. J. 1974. The laboratory mouse as a mammalian model for the genetics of growth. First World Congr. on Genetics Applied to Livestock Prod., Madrid, Spain 1:467.

Eisen, E. J. and B. S. Durrant. 1980. Genetic and maternal environmental factors influencing litter size and reproductive efficiency in mice. J. Anim. Sci. 50:428.

Eisen, E. J. and J. M. Leatherwood. 1978. Adipose cellularity and body composition in polygenic obese mice as influenced by preweaning nutrition. J. Nutr. 108:1652.

Eisen, E. J., J. Nagai, H. Bakker and J. F. Hayes. 1980. Effect of litter size at birth on lactation in mice. J. Anim. Sci. 50:680.

Eisen, E. J., W. R. Williams and J. F. Hayes. 1977. Effects of early pregnancy on postnatal maternal performance of mice. Anim. Prod. 25:1.

Falconer, D. S. 1973. Replicated selection for body weight in mice. Genet. Res. 22:291.

Harvey, W. R. 1975. Least-squares analysis of data with unequal subclass numbers. USDA, ARS H-4.

Hayes, J. F. and E. J. Eisen. 1979a. Environmental influences on body composition in mice selected for body weight. Theor. Appl. Genet. 55:209.

Hayes, J. F. and E. J. Eisen. 1979b. Direct genetic and postnatal maternal genetic effects on body

composition in mice selected for body weight. Can. J. Genet. Cytol. 21:43.

Hayes, J. F. and J. C. McCarthy. 1976. The effects of selection at different ages for high and low body weight on the patterns of fat deposition in mice. Genet. Res. 23:389.

Jagot, S. A., G. P. Webb, P. D. Rogers and J. W. T. Dickerson. 1980. Estimation of body fat content in normal, chemically-obese and obese hyperglycaemic mice. IRCS Med. Sci: Biochem; Biomed. Tech.; Endo. Sys.; Metab. and Nutr.; Physiol. 8:93.

La Salle, T. J. and J. M. White. 1975. Characterization of response to selection for growth and maternal ability in laboratory mice. J. Anim. Sci. 40:48.

Legates, J. E. 1972. The role of maternal effects in animal breeding. IV. Maternal effects in laboratory species. J. Anim. Sci. 35:1294.

Moore, R. W., E. J. Eisen and L. C. Ulberg. 1970. Prenatal and postnatal influences on growth in mice selected for body weight. Genetics 64:59.

Nagai, J. 1977. Evaluation of mouse lines for maternal performance and growth of offspring. Lab. Anim. Sci. 27:180.

Nagai, J., H. Bakker and E. J. Eisen. 1976. Partitioning average and heterotic components of direct and maternal genetic effects on growth in mice using crossfostering techniques. Genetics 84:113.

Rogers, P. and G. P. Webb. 1980. Estimation of body fat in normal and obese mice. Brit. J. Nutr. 43:83.

Skjervold, H. 1977. The effect of foetal litter size on milk yield: crossfostering experiments with mice. Z. Tierzüchtg. Züchtgsbiol. 94:67.

White, J. M., J. E. Legates and E. J. Eisen. 1968. Maternal effects among lines of mice selected for body weight. Genetics 60:395.

PAPER 20

The growth of mice selected for large and small size in relation to food intake and the efficiency of conversion.

Genet. Res. 38, 9-24. 1981

by

R.C. ROBERTS

Genet. Res., Camb. (1981), **38**, *pp.* 9–24
With 14 *text-figures*
Printed in Great Britain

The growth of mice selected for large and small size in relation to food intake and the efficiency of conversion

R. C. ROBERTS

ARC Unit of Animal Genetics, Institute of Animal Genetics, West Mains Road, Edinburgh EH9 3JN

(*Received* 7 *July* 1980 *and in revised form* 10 *December* 1980)

SUMMARY

Mice selected for large size show increases in both food intake and efficiency, and small mice show decreases in both. This is true whether the comparisons are made at the same age or at the same weight. Food intake and efficiency contributed more or less equally to the responses to selection for growth. Mice seem to regulate their food intake to a certain level of energy. On suspension of a period of food restriction, mice ate the same amount as others of the same strain that had not been restricted, and which were bigger. At the same time, they converted it more efficiently than the mice which had been full-fed throughout, because of a linear negative association between efficiency and body weight. Thus, following restriction, mice eat as much as bigger mice of the same age, and convert it as efficiently as younger mice of the same weight. The product of these two effects gives rise to rapid (compensatory) growth.

1. INTRODUCTION

If animals are selected to grow faster, it follows that they must do so by eating more, or by utilising the food more efficiently, or by some combination of the two. Typically, both voluntary food intake and efficiency of conversion change as positive correlated responses to selection for growth, as shown for a variety of species in a review by Yüksel (1979). The genetic correlation between the two component traits is usually small except in the pig, where the evidence is reasonably consistent that it is negative. In the mouse, the evidence is overwhelming that genetic variation exists both for voluntary food intake and for efficiency of conversion. Timon and Eisen (1970), Sutherland *et al.* (1970), Hayes & McCarthy (1976), Eisen, Bakker & Nagai (1977), Eisen & Bandy (1977) and McPhee *et al.* (1980) all agree that when mice are selected for increased weight gain, they both consume more food and convert it more efficiently, usually significantly so. A more detailed review of the mouse literature on this topic was provided by Roberts (1979). While the generality as stated can be defended, it is perhaps only fair to add that the magnitude of various changes in the two component traits is variable. The nature of the genetic variance in the system is not well understood.

Present address: ARC Animal Breeding Research Organisation, West Mains Road, Edinburgh EH9 3JQ.

The purpose of this paper is to report on the changes in food consumption and in efficiency found in the lines of mice selected for growth described by Falconer (1973). The material is unusual in that it comprises 6 large lines, 6 small lines and 6 unselected control lines. Variation among replicates within size groups is largely attributable to genetic drift, which allows us to examine how much of the variation in food intake and in efficiency may stem from this cause alone. Unfortunately, it has not been possible to examine all the replicates systematically, as measuring food intake of individual mice is demanding of facilities and labour. This paper summarises the main findings of a series of trials, usually conducted on a small scale, over several years. The results, for ease of summary, are presented graphically and the main conclusions are evident without statistical refinement. Some aspects of the data will be developed in further papers, but for present purposes, the emphasis will be on the coherent features of the separate studies.

Three lines of enquiry will be presented sequentially. The first is purely descriptive: how much food do large and small mice eat, and how well do they convert it? The only novel feature of this study is that the observations were continued until well after maximum body weight had been reached. The second experiment examines the regulation of food intake when alternative sources of energy are available. In particular, do any of the strains respond differentially to a ready source of energy, such as glucose? Finally, how do different strains react to the same amount of food, and what happens when they revert to their normal intake?

2. MATERIALS AND METHODS

All mice in these studies were housed and fed individually. The pelleted food was given to them in especially constructed baskets. Attempts were made periodically to estimate any waste, by keeping mice in cages without sawdust and bedding and separating out any food crumbs. Very little waste was actually observed, less (and usually much less) than 5% of the amount eaten, confirming the general impression that wastage was not a problem.

For the first study, mice were taken at 3 weeks of age from each of the large, control and small lines from the 17th generation of three replicates (replicates B, C and D) of Falconer's (1973) Q stocks. Each of the 9 lines provided 11 male mice. This trial was not terminated until the surviving mice were 75 weeks old. By that time, numbers were falling in some groups, but 81 of the 99 mice set up were still alive after one year. Body weights and food intake were recorded every week.

The next study came from the 27th generation of the Q stocks, and two of the above replicates (C and D) were represented, each with its large, control and small lines. The mice this time were all females, and each line was divided between two treatments, differing in the liquid that was provided. One was the usual tap water, the second a 5% glucose solution. The glucose was introduced in an attempt to manipulate any glucostatic mechanism of appetite control, since preliminary trials had indicated that mice would drink glucose solution preferen-

tially. A 5% solution was taken because this is isotonic with mouse body fluids, removing any complicating effect of tonicity on appetite. In this study, an attempt was made to monitor blood glucose level on the different treatments, but the measurement was unsatisfactory and the results are omitted from this paper. There were six mice from each line on each treatment and the trial was conducted between the ages of 3 and 8 weeks.

The last study to be described came from the 37th generation, replicates B and D being represented but the small lines being excluded. The two large lines and the two controls were fed *ad libitum*; in addition, a sample of each large line was fed on a daily basis exactly what the corresponding control line (at the corresponding age) had eaten the previous week. The trial began when the mice reached three weeks of age, and the intention was to restrict the large line to not less than 80% of its normal weight. By six weeks of age, it was becoming clear that the effects of the restriction were becoming more severe than this, so full feeding was immediately resumed. The mice were killed at 26 weeks of age and the gonadal fat was excised and weighed. However, no difference in carcass composition was detectable by then, and the results are omitted from this paper. There were 10 male mice per group in this study.

3. RESULTS

Growth, food intake and efficiency

Mean growth curves of the mice used in the first study are shown in Fig. 1. Selection had been for weight at 6 weeks of age but as expected, the weights were different at all ages. Mice reach about 80% of their mature weight by 10 weeks of age. All sizes gradually gain weight until they are at least a year old, presumably in part through the accumulation of fat. The three large lines are very similar in weight at all ages; the control and small lines show rather more divergence among themselves, but the nine lines differentiate clearly into three size groups.

For simplicity, food consumption is shown (Fig. 2) as the average of the three lines within a size group. There is no doubt that large mice eat more than small ones, and it would be astonishing if they did not. This is particularly noticeable at or around the age of selection, at 6 weeks, after which age the difference becomes less because the large mice reduce their intake. Food intake per unit of body weight is summarized in Fig. 3, with more detail in Table 1. Avoiding consideration of what exact power of weight may be the most appropriate, we may conclude the following. When mice are weaned at 3 weeks of age, large and small mice eat the same amount per unit body weight, but thereafter, this measure of intake declines differentially between the two. Small mice when fully grown eat their own weight of food per week, whereas large mice eat only three-quarters of their own weight. The control lines are very similar to the large lines (on this measure) until around 6 weeks of age, and thereafter adopt a level intermediate between large and small.

Large mice eat more than the controls, and small mice less, not only at the same age but also at the same weight (Table 2). Only a limited range of overlapping

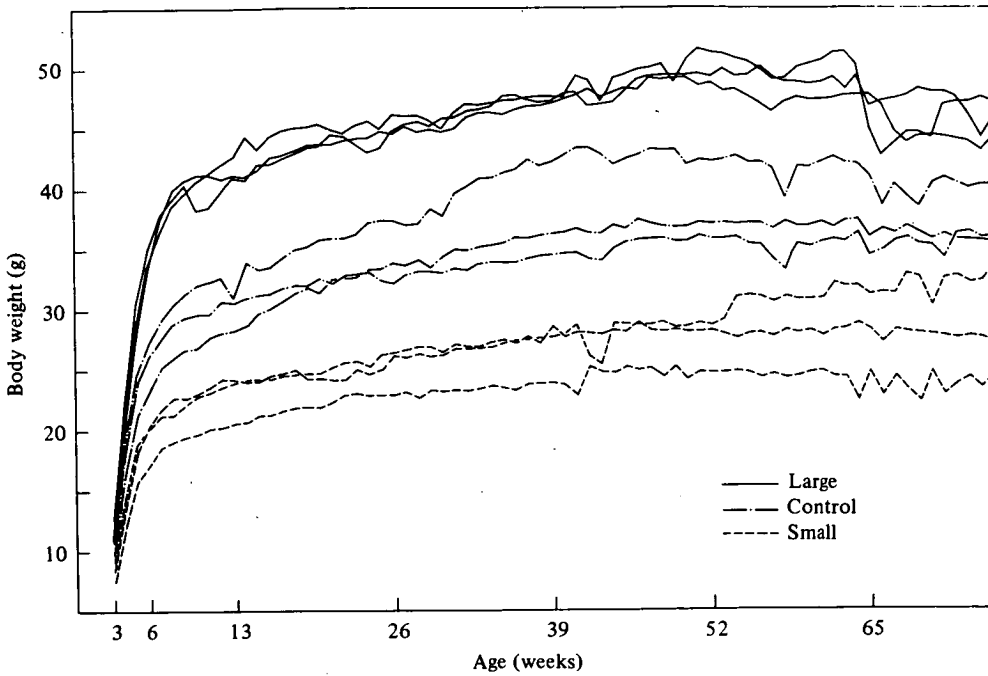


Fig. 1. Body weight in grammes of three replicates each of large (L), control (C) and small (S) mice, from 3 to 75 weeks of age.

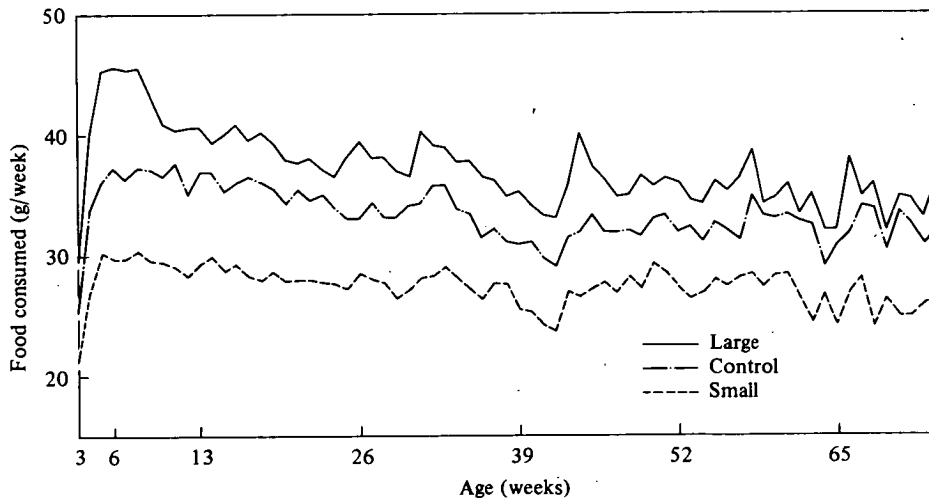


Fig. 2. Mean food consumption, in grammes per week, of the 3 replicates in each size group. L, C and S as in Fig. 1.

weights is available. Table 2 was constructed by determining the age at which each line reached 15 g and 20 g, and noting its food intake at those ages. The means for each size group are tabulated.

Efficiency of conversion was expressed as weight gain divided by food intake, a measure which declines to zero when growth stops. But over the growing period,

until about 8 weeks of age, there is no doubt that large mice are more efficient than the controls, and not so clearly, that the small mice are less efficient (Fig. 4). This is presumably related to the relatively higher maintenance requirement of the small lines, which is obvious at mature sizes (Fig. 3), as they require more food to maintain one unit of body weight. This is conventionally explained in terms of a higher surface area to mass ratio, with its implications for thermoregulation and its energetic cost.

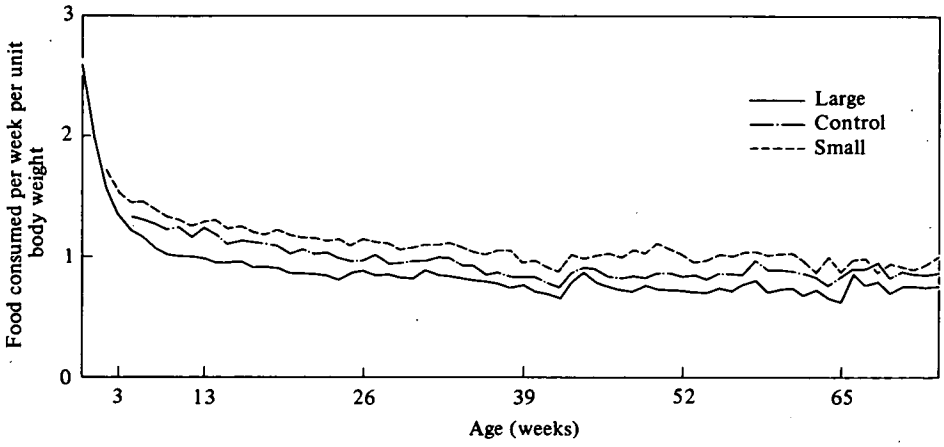


Fig. 3. Mean food consumption in grammes per week per unit body weight.

Table 1. *Food consumed per week per unit body weight at various ages. Averages for* 3 *replicates each of large* (*L*), *control* (*C*) *and small* (*S*) *mice*

	Age in weeks							
Lines	3	4	5	6	9	12	26	52
L	2·58	1·98	1·58	1·35	1·07	1·00	0·88	0·72
C	2·44	1·94	1·56	1·45	1·27	1·15	0·96	0·83
S	2·59	2·03	1·72	1·54	1·39	1·25	1·14	1·01

Table 2. *Average amount of food consumed by large, control and small mice at the same body weight*

	At 15 g		At 20 g	
Lines	Av. age (wk)	Food (g)	Av. age (wk)	Food (g)
L	3·2	33	3·8	37
C	3·6	31	4·4	34
S	4·3	28	7·1	30

We thus see that the effect of selection for growth has been to change both food intake and efficiency correspondingly. Since weight gain is the product of these two effects, we can compare the relative magnitude of their contributions to the change in weight. Table 3 shows the percentage changes of the large and small lines from their corresponding control line – each row indicating what

might have been observed in an unreplicated experiment. This, however, pays too much attention to the historical origins of large–control–small sets; genetic drift over generations means that the LB large line (for instance) has little more in common with the CB control line than it has with either of the other two controls. The best estimate of the relative contributions of food intake and efficiency to differences in body weight thus comes from the average of the three replicates, which shows the effects to be roughly equal. That one large line should have all of its response apparently attributable to improved efficiency, and one small line to reduced food intake, can be dismissed as accidents of drift, emphasizing that general conclusions cannot be drawn from single comparisons when lines are subject to drift.

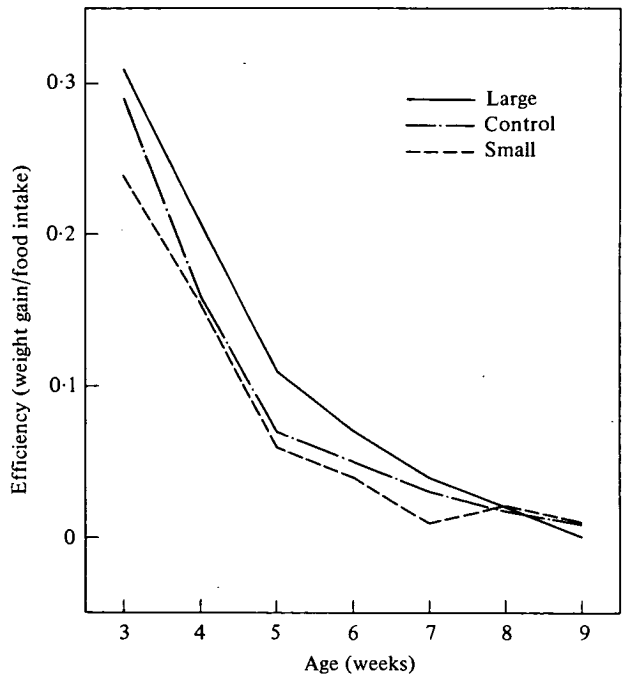


Fig. 4. Mean efficiency (weight gain/food intake) of the 3 replicates in each size group. L, C and S as in Fig. 1.

Table 3. *Percentage changes, relative to control, in voluntary food intake and gross efficiency (ages: 3 to 6 weeks)*

	Large lines		Small lines	
Rep.	Food intake	Efficiency	Food intake	Efficiency
B	+4	+46	−19	−19
C	+27	+28	−17	−1
D	+34	+30	−16	−14
Average	+22	+35	−17	−11

The conclusion from this study is therefore that selection for body weight changes both food intake and efficiency in the appropriate direction. The combination of genetic parameters in the two component traits is such that both can contribute more or less equally to the response.

Regulation of intake

The next study attempted to examine further the changes in voluntary food intake. Particularly, was the growth of large mice limited by their inability to ingest more food, or conversely, could their voluntary intakes be reduced by an alternative method of meeting their satiety requirements? We cannot enter here into detailed discussion of appetite control, but one simple theory (as described in standard biological text books) was amenable to test. This is the glucostatic

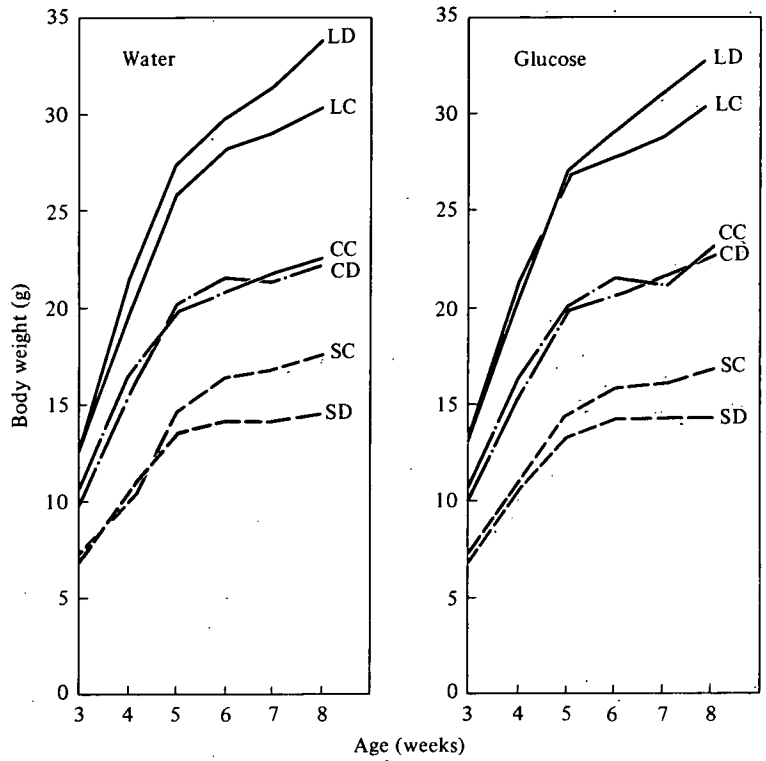


Fig. 5. Body weights (g) between 3 and 8 weeks of 2 replicates each of L, C and S mice. On left, when given ordinary water; on right, when water replaced by glucose solution.

theory of short-term appetite control. In simplified form, the theory states that the 'satiety centres', the ventromedial nuclei of the hypothalamus, respond to an elevated blood glucose level, and the animal stops eating until blood glucose drops again. The experiment asked a simple question: in view of the known predilection of mice to drink glucose solution in preference to ordinary drinking water, would this extra glucose in readily available form satisfy the satiety centre earlier, reduce food intake, and thus reduce growth? Or would it in fact provide extra energy to meet the energetic cost of protein synthesis (see Webster, 1977, for discussion) and thus allow the animals to grow more? To test this, each line was divided between two treatments, one being given ordinary tap water and the other a 5% glucose solution.

The glucose solution had no effect on weight gains – growth between 3 and 8 weeks was identical on the two regimes (Fig. 5). There were, however, obvious differences between lines in their glucose intake. The two large lines drank about $2\frac{1}{2}$ times as much glucose solution as they did of water, while the two small lines showed only a slight increase (Fig. 6). The interpretation is complicated by the behaviour of the two control lines, which went one each way. This division suggested that the control lines might have been fixed for different alleles controlling glucose preference, possibly alleles at a single locus. To test this, further samples from the six control lines maintained in the laboratory (Falconer, 1973) were measured for glucose solution intake from 6 to 12 weeks of age. The weekly intakes for the C and D control lines (which had shown a difference in Fig. 6) are

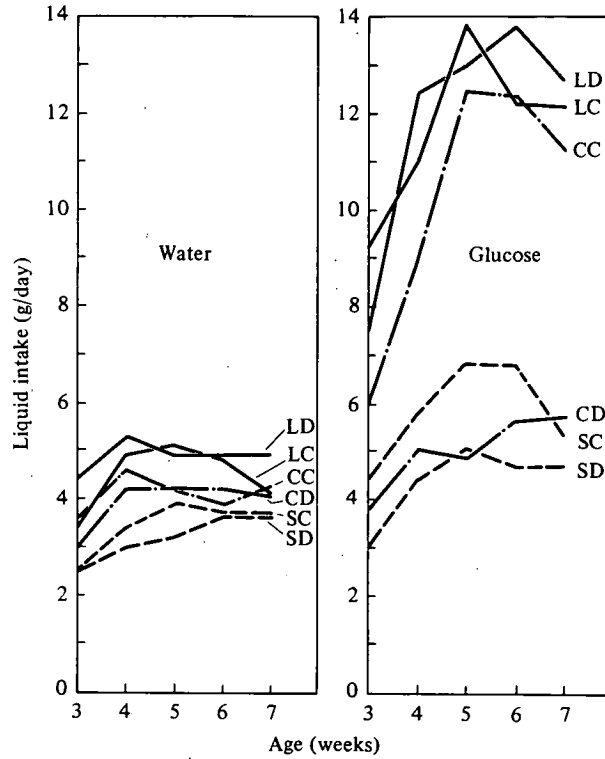


Fig. 6. Mean daily liquid intake (g) of water (on left) and glucose solution (on right) of L, C and S mice.

shown in Fig. 7. There is no hint of discontinuities in the distributions within lines, and unlike the previous sample, the mean intakes are not very different. The variation between individual mice was immense, and we can only suppose that the unfortunate difference between the controls in Fig. 6. was an accident of sampling small numbers. Given this, we shall treat the samples in Fig. 6 merely as two groups – glucose preferrers and non-preferrers.

The intake of solid food was reduced for the three lines that drank glucose solution copiously, compared to their intakes with normal drinking water, while

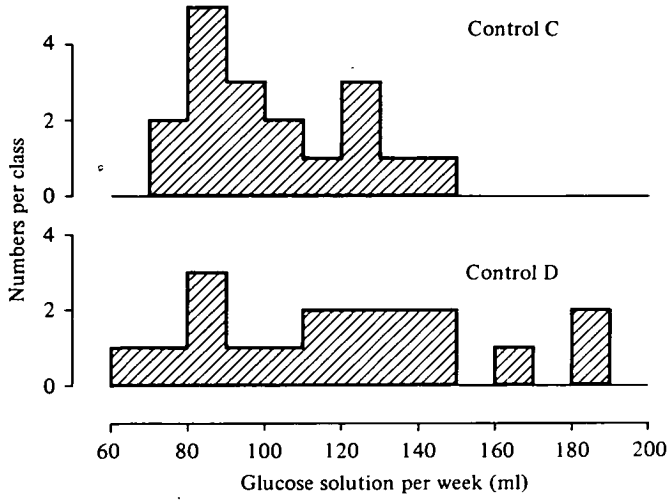


Fig. 7. Distributions of glucose intake within each of the two control strains, for the sixth week on trial.

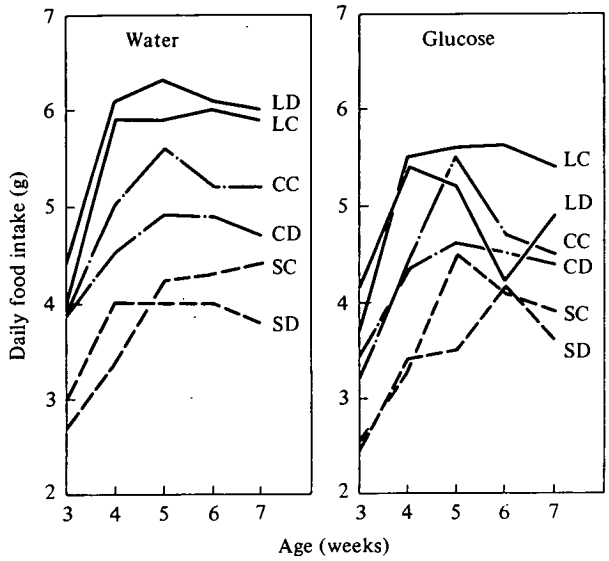


Fig. 8. Daily food intake (g) of mice when given water (on left) or glucose solution (on right).

that of the other three lines was affected much less (Fig. 8). The next step was to compare the total energetic intake from the two sources. The energetic value of glucose is available from standard text books, while that of our cubed diet was obtained by bomb calorimetry. The total energetic intake of mice of different strains was very similar whether they had water or glucose solution (Fig. 9). It is noticeable, though, that weekly fluctuations were greater among the three lines that drank glucose solution, as if they were having trouble with their glucostatic mechanism, but the overall impression is that of successful regulation of intake to

a given energy level. The regulation is too precise to be dismissed as a mere consequence of drinking so much glucose solution that they could not eat solid food.

The conclusion from this study was that whereas large mice may have a predilection for glucose solution, this does not induce satiety nor does it allow them to increase their total energetic intake. This suggests that the changes in appetite brought about by selection for increased growth are mediated, at least in the short term, through a mechanism related to total energetic requirements.

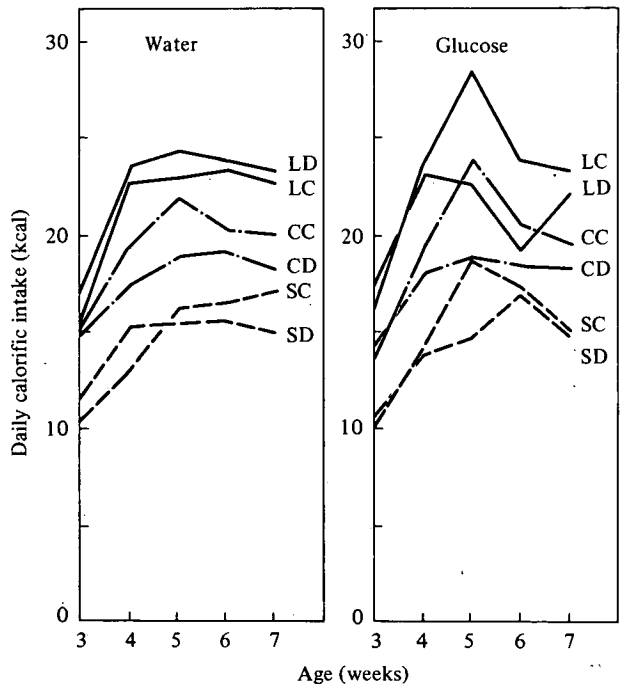


Fig. 9. Total daily calorific intake of mice when given water (on left) or glucose solution (on right).

Food restriction and compensatory growth

The final study involved restriction of food intake. If, as shown earlier, the large mice are more efficient than the controls, then on the same amount of food they ought to grow more. When this was tested, however, the large mice in fact gained less weight than the controls (Fig. 10) on the same amount of food, between the ages of 3 and 6 weeks. The amount fed to the large mice was the voluntary intake of the controls. This was barely sufficient to cope with their maintenance requirement, and at this level of feeding, the large mice were less efficient than the controls, unlike the situation under *ad libitum*. This reflects the complicated relationship between level of intake and efficiency, well-known to animal nutritionists. The result differs also from that reported by Stanier and Mount (1972). In a similar study, they fed both large and control mice the same amount of food

– 4 g per day. In their study, the large mice still grew better than the controls, though the differences between the two on *ad libitum* feeding were not fully realized under restriction. However, other differences between the two studies apart, the restriction imposed by Stanier and Mount was somewhat less severe for the large mice than the one used in this study, especially at younger ages, and that alone may be sufficient to explain the difference between the two results. Nevertheless, the finding that the large mice in this study gained less weight than the controls was unexpected.

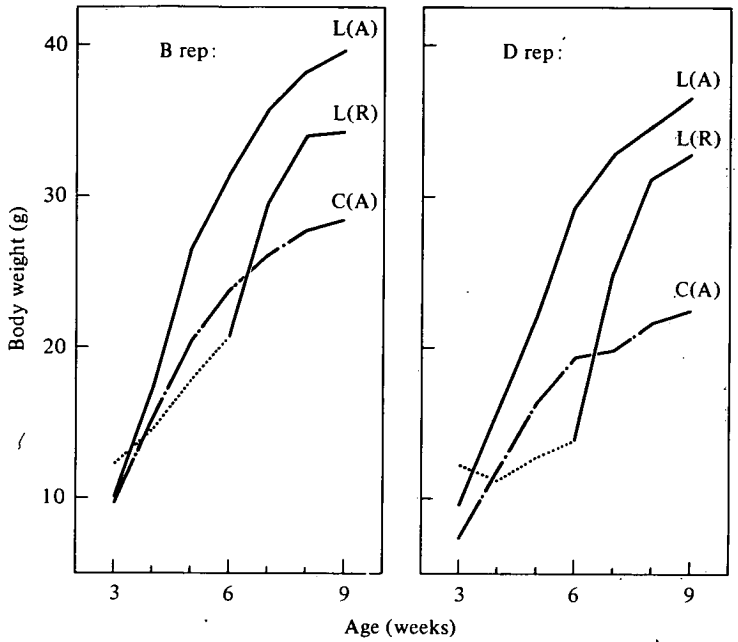


Fig. 10. Body weights (g) between 3 and 9 weeks of age in 2 replicates each of L and C lines. L(A) and C(A) were both fed *ad libitum*; L(R) was restricted to the food intake of the corresponding control, between 3 and 6 weeks of age.

At six weeks, the food restriction was terminated, and all mice were fed *ad libitum*. The two lines that had been restricted now showed the classical compensatory growth. Later weights are not shown in the interest of curtailing the graphs, but by 20 weeks of age, the restricted large line in the D replicate had caught up with the one fed *ad libitum* throughout. The B replicate seemed to settle down about 5 g less, but whether this was a difference between samples or a permanent effect of earlier restriction, we cannot tell.

The next question is how do animals achieve compensatory growth? Taking food intake first, the astonishing finding is that the very first week after being taken off restriction, the mice ate at least as much as those which had been full-fed (Fig. 11) and which were 60–100 % heavier than they were. There may be an element of chance in such precise regulation of appetite, but using words loosely, the system behaves *as if* voluntary food intake were simply a function of age.

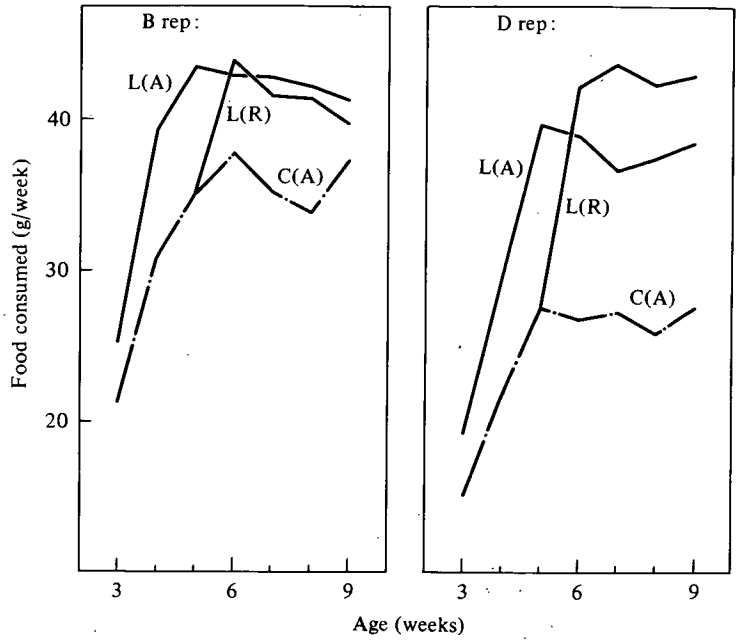


Fig. 11. Food consumption (grammes per week) of the lines shown in Fig. 9.

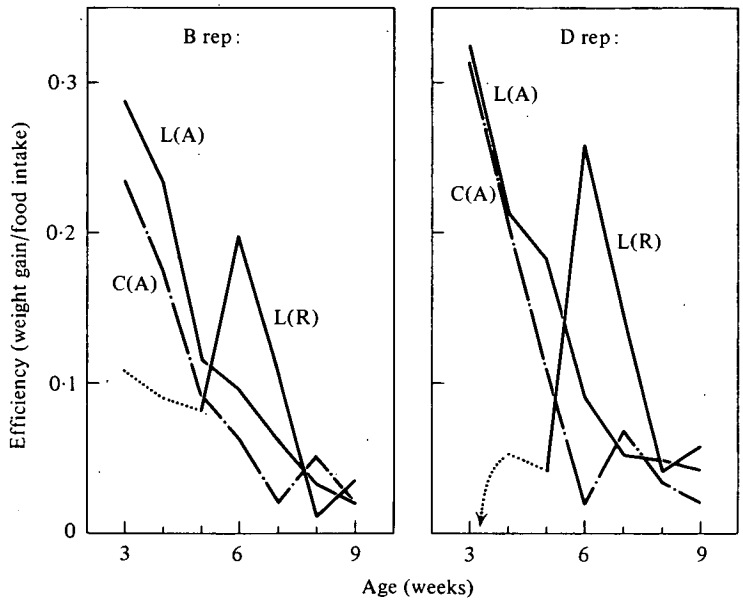


Fig. 12. Efficiency of conversion of the lines shown in Fig. 9.

Mice of the large lines, between 6 and 7 weeks of age, seem to eat a similar amount of food irrespective of whether they are 15 g or 30 g at the time.

Next, the lines that had been on restriction, as soon as they were taken off, show a large improvement in efficiency (Fig. 12), which then declines roughly in parallel to that of the full-fed mice. This spike in efficiency can be interpreted from the

following: the mean efficiency of the two large lines was plotted against their mean body weight at the time over weekly intervals between the ages of 3 and 10 weeks (Fig. 13). We see that efficiency declines as a remarkably linear function of body weight. When we superimpose on this plot those mice which had been restricted (marked as stars in Fig. 13) we see that they adopt the efficiency exactly appropriate to their weight at the time irrespective of their age.

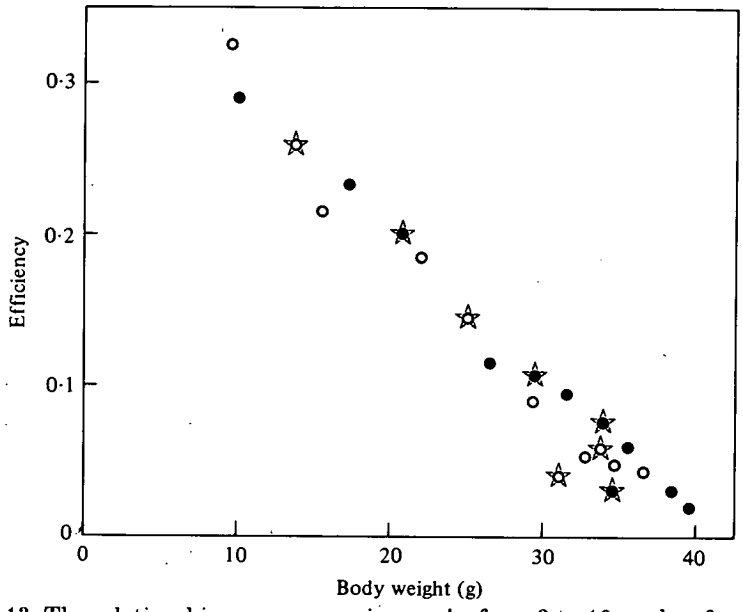


Fig. 13. The relationship, over successive weeks from 3 to 10 weeks of age, between efficiency and body weight at the time, for the two large lines fed *ad libitum*. Stars show the weekly efficiencies following restriction i.e. from 6 to 9 weeks.

It thus seems that, following a period of food restriction, mice have a bigger appetite than they normally would at that weight, and a higher efficiency than they normally would at that age. Because they are smaller and older, they benefit on both accounts, and the product of these two benefits is the rapid gain in weight known as compensatory growth.

In conclusion, one important consequence of selection should be noted. Selected large mice are more efficient than unselected controls at all body weights (Fig. 14), consonant with the earlier finding that they are more efficient even age for age, despite being bigger. The hint in Fig 14. that Large and Control mice differ in slope was not substantiated by a formal test.

4. DISCUSSION

There is no doubt that selection for large size in the Edinburgh lines of mice led to an increase both in food intake and in efficiency, and that this is true whether the comparisons are made at the same age or at the same weight. The small mice, correspondingly, show decreases in intake and efficiency. What is

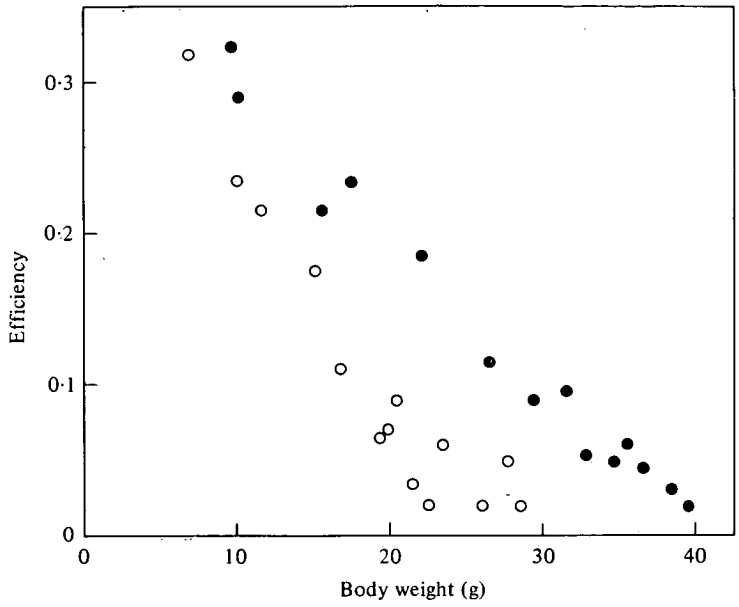


Fig. 14. The relationship between efficiency and body weight at the time, for the large and control lines, replicates pooled in both cases. Large, ●; control, ○.

more difficult to decide is to what extent the changes in the two component traits may be related. Other things being equal, an increase in food intake would be expected, at least over some range of intake, to lead to higher efficiency. If we subtract the maintenance requirement from a given intake, then what is left over is available for growth. Animals eating little more than their maintenance requirement will therefore not grow, and will be inefficient; those eating more will grow, and be more efficient. As we saw from the food restriction results, the amount eaten by the control lines was little above the maintenance requirement of large lines which explains why the large lines were very inefficient on the same feeding level as the controls.

However, the qualification for the mechanical relationship, stated above, was 'other things being equal', and they may not be. For a start, as we saw from the first study, small mice have a relatively higher maintenance requirement per unit body weight, seen clearly when all mice have stopped growing. This need not be wholly attributable to the higher surface area to mass ratio. Maintenance requirement covers not only thermoregulatory aspects but also protein turnover. Priestley and Robertson (1973), working with these strains, suggested that protein turnover was indeed slower in large mice than in small mice; small mice in fact had a higher rate of protein synthesis, but this was outweighed by even more rapid protein degradation. One of the effects of selection may thus have operated on protein turnover, which may explain why large mice are more efficient than the small mice when compared at the same body weight, despite the strong negative association (within a size group) between weight and efficiency.

It seems therefore that selection for weight is not simply selection for increased

appetite, and that appetite may in fact be some consequence of metabolic changes brought about by the selection. Radcliffe and Webster (1976), working on the rat, suggested that appetite control is linked to the animal's impetus for protein deposition, and that the retention of lipid and energy is of no consequence for appetite control. If we may suppose that this 'impetus' is the genetic regulation of body size (though that statement is no more precise) it then follows that the increased food intake of our large mice is a consequence of their more rapid growth. This may explain the precision of appetite control found here both in the glucose and in the restriction studies. It may also explain the curious hump in the food intake of the large lines around 6 weeks of age, following a period of very rapid growth. As growth slowed down the metabolic demands were reduced sufficiently to cut back their voluntary food intake to a lower level.

There are indications from studies on domestic livestock that the results reported here on the effects of restriction may have some generality. Saubidet and Verde (1976) reported that Aberdeen Angus steers, subjected to varying levels of restriction, all had very similar intakes once the restriction was removed. They note specifically that steers of the same age had very similar intakes, irrespective of their weights at the time. They also note that the restricted animals had a lower maintenance requirement because of their lower body weights, thus implying higher efficiency once restriction was removed. Somewhat similar effects were also found in sheep (Allden, 1968, 1970), but with the qualification that this depends on the ages over which the animals are restricted.

The implications of the mouse studies reported in this paper are that growth, food intake and efficiency are all inter-related whereby no single member can be fully understood independently of the other members. Ideally, carcass composition should also be included. Not only would this give a more complete picture of the input–output system, but it might also indicate genetic factors involved in the partitioning of metabolites to various destinations. It seems that selection for growth may have operated, at least in part, at the level of this partitioning.

REFERENCES

ALLDEN, W. G. (1968). Undernutrition of the Merino sheep and its sequelae. I. The growth and development of lambs following prolonged periods of nutritional stress. *Australian Journal of Agricultural Research* **19**, 621–638.

ALLDEN, W. G. (1970). The effects of nutritional deprivation on the subsequent productivity of sheep and cattle. *Nutrition Abstracts and Reviews* **40**, 1167–1184.

EISEN, E. J., BAKKER, H. & NAGAI, J. (1977). Body composition and energetic efficiency in two lines of mice selected for rapid growth rate and their F_1 crosses. *Theoretical and Applied Genetics* **49**, 21–34.

EISEN, E. J. & BANDY, T. (1977). Correlated responses in growth and body composition of replicated single-trait and index selected lines of mice. *Theoretical and Applied Genetics* **49**, 133–134.

FALCONER, D. S. (1973). Replicated selection for body weight in mice. *Genetical Research* **22**, 291–321.

HAYES, J. F. & MCCARTHY, J. C. (1976). The effects of selection at different ages for high and low body weight on the pattern of fat deposition in mice. *Genetical Research* **27**, 389–403.

McPHEE, C. P., TRAPPETT, P. C., NEILL, A. R. & DUNCALFE, F. (1980). Changes in growth, appetite, food conversion efficiency and body composition in mice selected for high post-weaning weight gain on restricted feeding. *Theoretical and Applied Genetics* **57**, 49–56.

PRIESTLEY, G. C. & ROBERTSON, M. S. M. (1973). Protein and nucleic acid metabolism in organs from mice selected for larger and smaller body size. *Genetical Research* **22**, 255–278.

RADCLIFFE, J. D. & WEBSTER, A. J. F. (1976). Regulation of food intake during growth in fatty and lean female Zucker rats given diets of different protein content. *British Journal of Nutrition*, **36**, 457–469.

ROBERTS, R. C. (1979). Side effects of selection for growth in laboratory animals. *Livestock Production Science* **6**, 93–104.

SAUBIDET, C. L. & VERDE, L. S. (1976). Relationship between liveweight, age and dry matter intake for beef cattle after different levels of food restriction. *Animal Production* **22**, 61–69.

STANIER, M. W. & MOUNT, L. E. (1972). Growth rate, food intake and body composition before and after weaning in strains of mice selected for mature body weight. *British Journal of Nutrition* **28**, 307–325.

SUTHERLAND, T. M., BIONDINI, P. E., HAVERLAND, L. H., PETTUS, D. & OWEN, W. B. (1970). Selection for rate of gain, appetite and efficiency of feed utilization in mice. *Journal of Animal Science* **31**, 1049–1057.

TIMON, V. M. & EISEN, E. J. (1970). Comparisons of *ad libitum* and restricted feeding of mice selected and unselected for postweaning gain. I. Growth, feed consumption and feed efficiency. *Genetics* **64**, 41–57.

WEBSTER, A. J. F. (1977). Selection for leanness and the energetic efficiency of growth in meat animals. *Proceedings of the Nutrition Society* **36**, 53–59.

YÜKSEL, E. (1979). Genetic aspects of the efficiency of food utilization in some farm and laboratory animals. *Animal Breeding Abstracts* **47**, 499–504.

PAPER 21.

A case of polydactyly with multiple thresholds in the mouse.

Proc. R. Soc. Lond. B. 191, 427-444
1975

by

R.C. ROBERTS & NANCY R. MENDELL

Proc. R. Soc. Lond. B. **191**, 427–444 (1975)
Printed in Great Britain

A case of polydactyly with multiple thresholds in the mouse

BY R. C. ROBERTS† AND NANCY R. MENDELL‡§

†*Agricultural Research Council, Unit of Animal Genetics,*
‡*Institute of Animal Genetics, West Mains Road,*
Edinburgh EH9 3JN

(*Communicated by D. S. Falconer, F.R.S.* –
Received 27 *March* 1975)

A spontaneous occurrence of polydactyly in mice failed to comply with any acceptable segregation ratio, but its genetic basis was established by its rapid response to selection. Both the incidence and the severity of abnormality increased. Initially manifested as one extra digit on one hind foot only, the polydactyly spread to both hind feet, and the number of supernumerary digits rose. Subsequently, the fore feet became affected as well, and in those cases, the hind legs became deformed in a manner reminiscent of the *luxate* mutant.

The genetic interpretation of the data was compatible with a model of liability to the abnormality based on a continuously distributed underlying variable, with two thresholds, one corresponding to hind foot polydactyly alone and a higher threshold which, when exceeded, made the mice polydactylous on the fore feet as well. Further analyses indicated that the threshold for the hind foot condition was in fact a cluster of thresholds, corresponding to increasing digit numbers. The data revealed complex interactions between systemic and local influences on digit number, with implications for the gene control of limb development.

INTRODUCTION

A comprehensive list of the known mutant genes in the mouse (see, for instance, recent issues of *Mouse News Letter*) would include at least twenty separate loci at which a mutant, among other possible effects, alters the number of digits on the feet. While many of these mutants affect the number of digits directionally, though perhaps irregularly, others seem to lead to a developmental instability which may result in either an increase or a decrease in number. However, the details of these mutants are of less concern to us here than the general picture which they present. It appears that the relatively invariant pentadactyly, which features a broad band of the taxonomic spectrum, is under the control of many genes in the mouse. It is reasonable to suppose that the twenty or more loci already known to be involved are but a proportion of the total, and further,

§ Present address: Department of Microbiology and Immunology, Duke University Medical Center, Durham, North Carolina 27710, U.S.A.

that underlying genetic variation in digit number may be present to an extent perhaps not suggested by the normal developmental stability of the character.

An indication of the general nature of the genetic system affecting digit number in the mouse was provided by Fisher as long ago as 1950. He was working on Holt's (1945) *polydactyly* (*py*), a recessive mutant which at first had been poorly manifested, with incomplete penetrance. Fisher had been able to select for improved manifestation and penetrance. Even so, Fisher stated that homozygotes for the mutant were not unconditionally polydactylous, and he found it necessary to invoke at least three modifying genes which reduced the expression of polydactyly. He was able to assign one of the modifiers to a linkage group; the other two remained unlocated, but it was claimed that they had been sufficiently isolated to make systematic tests possible. Fisher referred to polydactyly in mice as 'a situation of more than ordinary complexity', but considered that his results had 'gone so far as to clear up the situation to a very large extent'.

Some further clearing up, however, became necessary. Bodmer (1960), working on Fisher's stock, reported that the manifestation of polydactyly was increased (unlike Fisher's decrease) by two mutant genes, *pallid* and *fidget*, known primarily for the effects suggested by their names. Further, Bodmer reported that the modifying effect of the genetic background of the stock, including in particular *pallid* and *fidget*, was sufficient to yield some polydactylous mice even in the absence of the gene *py*, through which the condition had first become manifest.

Fisher and Bodmer, between them, thus found at least five loci which modified the expression of one polydactylous mutant, even to the extent of over-riding either homozygote at the so-called 'major' locus. When this is the case, the distinction between major and minor genes becomes somewhat eroded, though this was not clearly the case for *polydactyly* until Fisher and Bodmer had established the modifying strength of the genetic background. The number of loci that these two workers found which affected the expression of *py* is not necessarily exhaustive; indeed, it is unlikely to be. It is probable that other mutants affecting digit number can also be modified, and it appears certain that the twenty-odd loci known to contribute to the trait is, as suggested earlier, a gross underestimate of the total number of loci involved in the developmental system.

The broad outlines of the genetic control of digit number in the mouse emerging from these considerations strongly resemble the model proposed by Wright (1934), in his classical paper on polydactyly in guinea pigs. Wright regarded the genetic variation as continuous on some underlying scale, with a truncation point, which he called a 'threshold', marking the distinction between normality and polydactyly on the phenotypic scale. In most populations, the distance between the population mean and the threshold is such that few, if any, individuals lie beyond the threshold. Only when the gap is reduced sufficiently is the genetic variation on the underlying scale brought into play, and the most liable individuals then exceed the threshold and become polydactylous. The concept of 'liability'

and its inheritance was developed by Falconer (1965), in connection with certain human diseases. The essence of Falconer's treatment was to render threshold characters more amenable to analysis by the standard methods of quantitative genetics, which the genetic basis of such characters seems to demand.

While the ultimate objective must be to understand the biochemical and developmental pathways controlling digit number, where the action of every gene can be pinpointed, this ideal for the present seems remote. Detailed studies on mutants affecting digit number, as exemplified by Johnson (1969*a*, *b*) and Forsthoefel (1962), contribute greatly to the understanding of interrelationships between various parts of the developmental system controlling digit number. In this paper, however, we shall not be concerned with a single gene effects, because no single gene was isolated. We shall describe a polydactylous condition where the genetic variation had to be treated as continuous, and the results will be presented in terms of the appropriate methodology.

Materials and methods

The polydactylous condition to be described in this paper arose spontaneously in the fifth generation of a stock of mice (CQ) selected for large size, described by Roberts (1967). The possibility that polydactyly is in any way connected with selection for large size may be entirely discounted. The CQ stock at the time was not composed of particularly large mice, and a further fifteen generations of selection for increased body size did not yield any further cases of polydactyly. The original mating produced a male in its first litter that showed preaxial polydactyly (extra 'big toe') on its right hind foot. The mating was continued and over seven litters, it yielded 5 polydactylous offspring out of a total of 53. It was apparent at an early stage that the expression of polydactyly was variable. The position of the extra digit was not constant, and its size varied considerably. The usual procedures of backcrossing, intercrossing and outcrossing were employed in an attempt to establish the polydactyly as a new mutant. Without going into details, it became clear that the new polydactyly was not going to behave as a single-gene condition. It appeared sporadically throughout the stock, mostly as if it were a dominant of low penetrance, which is a singularly unhelpful genetic statement. But occasionally, it also appeared among the progeny of parents which were both unaffected. A large and messy pedigree chart was constructed. Suffice it to say that the polydactyly appeared only among the lineal descendants of one of the affected mice from the original mating; of the unaffected progeny of that mating which were tested, none ever gave rise to any abnormality unless mated to an affected sib. Even matings between two affected parents yielded rather few polydactylous offspring. One conclusion had already emerged – that a single gene hypothesis was proving unfruitful, even to the extent that the condition could not be described adequately as either dominant or recessive.

Other findings from these matings had prognostic features. The first was the

variability in the expression of the polydactyly. Secondly, one of the affected males from the original mating, when mated to a normal sib, produced a bilaterally affected offspring. These findings contributed to the general untidiness of the situation, and the experimental approach clearly had to be modified.

Partly to accommodate these considerations, and partly to counteract the inbreeding that was accumulating, fifteen surviving polydactylous mice of breeding age were pair-mated with unrelated mice drawn from the ninth generation of the CQ stock, in which the condition had first arisen. The intention was to test anew for segregation, after selection for increased incidence of polydactyly. However, the results of the selection themselves became the main topic of study. In order to limit space requirements, within-family selection was practised, thereby increasing the effective number of the stock. Where possible, two offspring (one of each sex) were selected from each family, and mated with other selected animals at random, save only that common grandparents were avoided. Except where stated otherwise, ten pair matings were set up each generation.

The criteria for selection were slightly informal and variable. During the early generations of the selection programme, the main criterion was the total number of extra digits. This eventually was modified, as explained later. The results of the selection are presented in the next section, together with an analysis of the genetic control of polydactyly in this stock of mice.

Results

(*a*) *Selection for increased polydactyly*

The main results of the selection, as indicated by incidence of abnormalities, are shown in figure 1. Some further details of the early generations are summarized in table 1. Briefly, as the incidence of polydactyly rose in the stock, so did the number of extra digits and other manifestations of severity. In generation 1, bilaterally affected mice became common, whereas earlier they had occurred only sporadically. Also in that generation, it became clear that the incidence of polydactyly was rare unless at least one parent was affected. This was confirmed by the reversed selection in generation 2 (see table 1). But the most dramatic change was the extension of the abnormality to the fore feet (which had previously remained unaffected) coupled with deformities of the hind legs, phenotypically closely resembling the effects of the *luxate* gene. Extra digits on the fore feet were generally associated with hind leg deformity, though the correspondence was not exact. The 'luxated' males proved to be sterile, quite possibly because of their physical incapacity to mate. 'Luxated' females, on the other hand, were normally fertile, and from generation 5 onwards, most matings were between such females and the 'non-luxated' males with the most supernumerary digits on their hind feet.

All the offspring of generation 5 were polydactylous, and it appeared that the condition had become fixed in the stock. This, however, proved to be wrong,

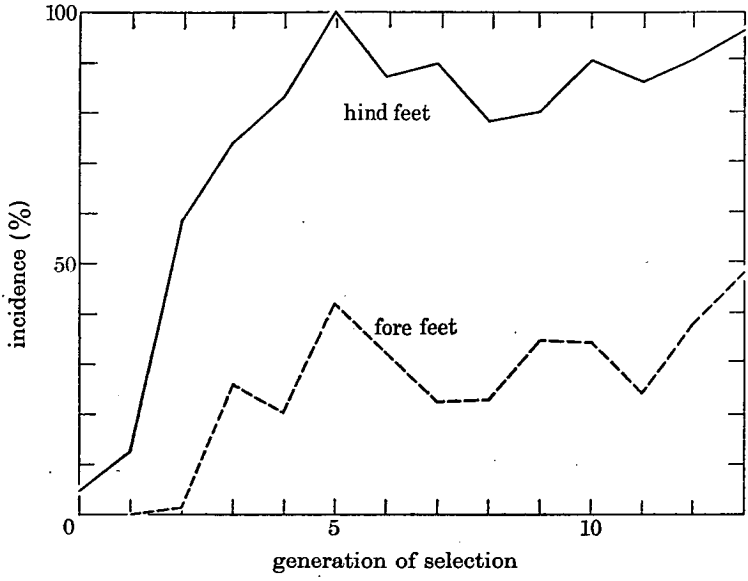


FIGURE 1. Incidence of polydactyly, on the hind and fore feet, at each generation of selection.

TABLE 1. SUMMARY OF RESULTS IN EARLY GENERATIONS OF SELECTION AND OF QUALITATIVE CHANGES IN MANIFESTATION

generation	no. of matings (no. fertile)	no. of offspring classified	no. of polydactylous offspring (incidence)	remarks
0	15 (14)	270	13 (0.048)	matings heterogeneous: 9 had no polydactylous offspring, others 1/24, 1/22, 1/17, 3/18, 7/13, respectively
1	16 (16)	278	35 (0.126)	6 matings had one polydactylous parent: 34/35 polydactylous offspring from these; 20/35 bilaterally affected; 11/35 had 7 digits on at least one hind foot
2	10 (8)	77	45 (0.584)	both parents polydactylous: one offspring had an extra digit on each fore foot, as well as on hind feet, with 'luxation' of hind legs
2R	10 (10)	76	2 (0.026)	reversed selection: parents unaffected (but had affected sibs)

and despite continued selection, up to 20% of all offspring in subsequent generations were unaffected. The incidence of fore foot abnormality wavered around one-third. A graph for the 'luxated' condition would follow closely the one for fore feet.

Among the affected offspring, there was every gradation of abnormality, from one extra digit on one hind foot upwards. There was also considerable variation between feet within affected mice. No case was found of fore-foot polydactyly without at least some deformity of a hind foot. There is a hint in figure 1 that the fluctuation in incidence between generations is similar for the hind and front feet, but the correspondence in pattern is not convincing.

The reversed selection (first tested in generation 2) was repeated in generations 4 and 5. Matings between unaffected individuals in these two generations yielded 17 and 13% polydactylous offspring, respectively. The corresponding incidence in generation 2 had been only 3% (see table 1). There is therefore some evidence that a general proclivity to polydactyly was building up even among unaffected members of the stock, but this was not tested further.

It remains an untestable presumption that the incidence of polydactyly would have increased further but for the sterility of the worst affected ('luxated') males. The potential fertility of these males should have been tested by artificial insemination, but the experiment was destined to be terminated before this procedure could be adopted.

The effect of selection was to yield a stock of mice in which the incidence of polydactyly was approximately 90%, accompanied in the worst cases by a skeletal abnormality of the hind limb.

(*b*) *Skeletal abnormalities of the hind limbs*

Approximately 30% of the mice from generation 5 onwards displayed abnormalities of the hind legs reminiscent of the effects of the *luxate* gene. In fact, the similarity is more than superficial. Figure 2 indicates the skeletal deformities found in a typically abnormal hind leg. The tibia is much reduced in size and is of irregular shape; the fibula is less affected but has become separated from the tibia. These changes, and the consequent displacements in articulation, closely resemble one of the *luxate* conditions described by Carter (1951). Carter illustrates five degrees of abnormality in order of ascending severity. The leg illustrated in figure 2 could be interchanged with the second of Carter's five without excessive misrepresentation. It is, however, only fair to add that unlike Carter's thorough and careful investigation of luxate, the abnormality described here was not studied in any detail, and any suggestion of an exact resemblance with the effects of the *luxate* gene is not warranted.

It is not clear why the hind-leg abnormality should have followed so closely the presence of fore-foot polydactyly. In a discussion of *luxate* and similar genes, Grüneberg (1964) points out that a combination of excess formations (polydactyly) in the foot with defects of the zeugopodium suggests a competition phenomenon

for a limited material as their cause. Following this suggestion, the association of fore-foot excesses with defects in the hind limbs indicates that such competition has to be systemic rather than local in origin. A superficial examination of the remainder of the skeleton indicated that the malformation of bones was not confined to the hind limbs, though such further abnormalities were less drastic on a visual and subjective appraisal.

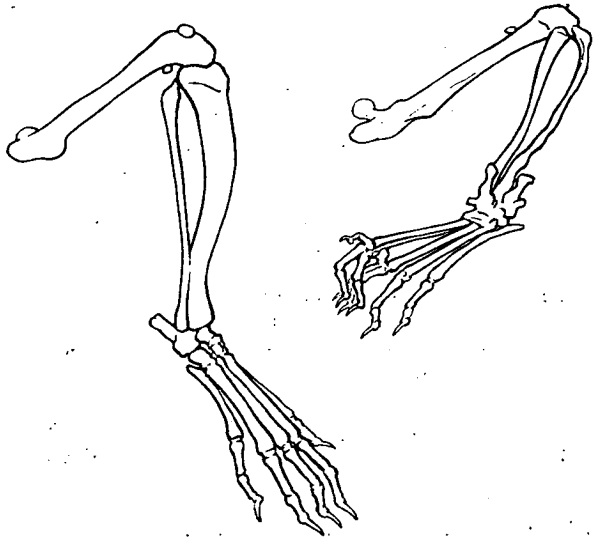


FIGURE 2. General nature of the skeletal abnormalities of the hind limbs, associated with fore-foot polydactyly. Normal hind limb on the left, for comparison.

Finally, in this section, it should be noted that the foot abnormality was not restricted to the mere presence of additional digits. The foot bones could be misshapen and disarticulated, and the foot illustrated in figure 2 could perfectly well have appeared on a mouse with no visible abnormality in the less distal part of the limb.

(c) *Modular units of abnormality*

As discussed earlier, the manifestation of the polydactyly varied, both with respect to the number of feet affected and the number of digits per affected foot. The initial appearance of each abnormality almost implied a step-wise progression of severity. From an unspectacular origin of one extra digit on one hind foot, selection generated a bilateral condition, to be followed by more than one extra digit per foot until finally, the abnormality extended to the front feet, accompanied in most cases by skeletal deformities of the hind legs. This pattern in the history of the stock suggested multiple thresholds on the phenotypic scale, in ascending order of severity. This section explores whether similar thresholds are equally distinguishable on a genetic scale.

(i) *Bilateral symmetry*

Superficially, there may be a suggestion that the bilateral condition is a more severe form of the abnormality than unilateral polydactyly, for two reasons. First, as soon as bilaterally affected animals became available for selection as parents, bilaterality of affected offspring became the rule rather than the exception.

TABLE 2. CLASSIFICATION OF OFFSPRING FROM GENERATIONS 2–13, INCLUSIVE, BY THE NUMBER OF EXTRA DIGITS ON THE TWO HIND FEET

no. of extra digits	no. of offspring classified by left foot	right foot
0	213	198
1	560	573
2	244	246
total	1017	1017

Expectations of various combinations if feet are independent:

		right 0	1	2	totals
left	0	41.47	120.01	51.52	213
	1	109.03	315.52	135.45	560
	2	47.50	137.47	59.03	244
	totals	198	573	246	1017

Expectations recalculated omitting 160 unaffected mice:

		right 0	1	2	
left	0	2.35†	35.44	15.21	53
	1	24.83	374.42	160.75	560
	2	10.82	163.14	70.04	244
	totals	38	573	246	857

(† unobservable)

Observed numbers in various classes:

		right 0	1	2	
left	0	160	40	13	213
	1	31	413	116	560
	2	7	120	117	244
	totals	198	573	246	1017

Secondly, the reversed selections in generations 2, 4 and 5 yielded a significantly higher proportion of unilaterally affected offspring than was found among the corresponding generations of forward selection. However, evidence of this nature is untrustworthy. Quite clearly, if the probability of extra digits is raised, then even a random allocation of extra digits between feet would increase the incidence of bilaterality, and vice versa. The correct question to ask therefore is whether

there is any tendency towards or against symmetry, i.e. does the incidence of bilaterality differ from the joint probability when the two feet are treated as independent manifestations of polydactyly.

Taking all animals from generation 2 onwards, when bilaterality became common, the classification of the number of extra digits on the hind feet is shown in table 2. The expectations for various combinations of left and right are calculated as if the two feet were independent, and these expectations are compared

TABLE 3. CLASSIFICATION OF ALL OFFSPRING FROM GENERATIONS 2–13, INCLUSIVE, BY THE NUMBER OF EXTRA DIGITS ON THE TWO FORE FEET

no. of extra digits	on left foot	on right foot
0	742	804
1	181	156
2	94	57
total	1017	1017

Expectations of various combinations if feet are independent:

		right			totals
		0	1	2	
left	0	586.60	113.82	41.58	742
	1	143.09	27.76	10.15	181
	2	74.31	14.42	5.27	94
	totals	804	156	57	1017

Expectations recalculated omitting 727 unaffected mice:

		right			
		0	1	2	
left	0	3.98†	8.07	2.95	15
	1	48.06	97.37	35.57	181
	2	24.96	50.56	18.48	94
	totals	77	156	57	290

(† unobservable)

Observed numbers in various classes:

		right			
		0	1	2	
left	0	727	11	4	742
	1	63	97	21	181
	2	14	48	32	94
	totals	804	156	57	1017

with the observed frequencies. Various χ^2 analyses can be done on these data, the essential comparison for now being that of the diagonal (symmetry) versus the off-diagonal (asymmetry). Even this comparison can be debated. There is a case, for instance, that the '00' class should be excluded, on the grounds that these are unaffected animals and are known to be qualitatively different – and, of

course, symmetrical. The expectations can be recalculated accordingly, based only on affected animals, with the hypothetical '00' class now becoming unobservable. But no matter how the comparison is done, there is unmistakably a significant excess of animals on the diagonal. Therefore, although asymmetry was by no means rare, there can be no doubt that among polydactylous animals, there was a distinct tendency towards symmetry, all the χ^2 values being significant well below the 0.1 % level. This presumably reflects some systemic influence on the development of the two hind limbs. But the extensive presence of asymmetry requires also a degree of local autonomy in the development of a limb. Indeed, the initial manifestation of unilateral polydactyly required such autonomy.

It is also seen from table 2 that, with regard to the hind feet, the mean number of extra digits was virtually identical on the left and right sides, with no suggestion of any differences in liability between the two. This, however, is not so with respect to the fore feet, as can be seen from table 3. Here the left fore foot is clearly more prone to polydactyly than the right foot. Although bilateral asymmetry is a feature of many organs, we had not particularly expected to find it reflected in the development of the fore limbs, especially as the hind limbs had tended towards symmetry. However, given this initial difference in mean liability between the two sides, the various combinations of excess digits on the two sides are much as expected. If we treat as unobservable the mice expected to show no fore-foot polydactyly, from the independent incidences of the two sides of mice affected on the fore feet, the distribution of the other classes shows no gross departures from expectation. The heterogeneity, it is true, is statistically significant, but by any absolute standard it must be judged unimportant. In particular, and quite unlike the hind feet, there seems to be little tendency of affected mice to pile up on the diagonal in table 3. Thus the incidences on the two fore feet, although unequal, behave as if they were independent.

In summary, the results from this section show a clear tendency towards the basic symmetry of manifestation on the hind feet, whereas the fore feet are basically unequal in incidence. We fail to suggest a plausible reason for the asymmetry of the fore feet.

(ii) *Fore foot polydactyly as an indicator of severity*

Phenotypically, there can be no doubt that extra digits on the fore feet represent a more severe form of polydactyly than the hind-foot condition alone. There are two reasons for this statement. First, fore-foot polydactyly did not appear without the hind feet being affected as well, whereas the converse was a frequent occurrence. Secondly, fore-foot polydactyly was strongly associated with other skeletal deformities, which typically were not apparent if the fore feet were normal.

The genetic analysis amply confirms this impression with regard to severity. All parents and offspring of generations 2 to 13, inclusive, were regrouped according to whether they were unaffected (O), affected on the hind feet only

(H) or on the fore feet as well (H + F). There were only four types of mating, and the numbers among two of those were small. Nevertheless, the results shown in table 4 are entirely consistent. Not only does the incidence of affected offspring rise according to the average severity of the parents, but also the incidence of the more severe form (H + F) rises in parallel, the statistical significance of the increase in both cases being beyond doubt.

Two conclusions emerge from this section. The first is that genetically, as well as phenotypically, fore-foot polydactyly is a more severe form of the abnormality than hind-foot polydactyly alone. The second conclusion derives from the close

TABLE 4. CLASSIFICATION OF OFFSPRING, WITHIN MATING TYPES, ACCORDING TO WHETHER THEY ARE FREE OF POLYDACTYLY (O), AFFECTED ON THE HIND FEET ONLY (H), OR AFFECTED ON THE FORE LIMBS AS WELL (H + F)

Data from same source as table 3, supplemented by information from reversed selections to increase the representation of the less common mating types.

mating type (♀ × ♂)	proportion of offspring with			total number of offspring
	O	H	(H + F)	
O × O	0.855	0.145	0	83
H × O	0.516	0.387	0.097	62
H × H	0.228	0.531	0.241	382
(H + F) × H	0.087	0.565	0.348	561
numbers per class	239	556	294	1088

correspondence between the incidences of the more- and less-severe forms, among different mating types, as seen in table 4. This argues strongly that the two forms are not separate genetic entities, but that they represent different areas from the one distribution. Some genetic correlation was of course already implied by the history of the selection programme. The data are compatible with Falconer's (1965) extension of Wright's (1934) original model of liability on an underlying scale, with two thresholds, one corresponding to hind-foot polydactyly alone and a higher threshold at which, when transcended, the condition extends to the fore feet. Thus, the threshold for the fore-feet manifestation cannot be exceeded without a hind foot being also affected.

(iii) *Variable number of supernumerary digits in parents*

Effect on incidence in offspring. Taking the simplest possible system of classification first, namely the total number of extra digits in both parents, the incidence of polydactyly among the offspring shows a marked increase as the number of digits in the parent increases (solid line in figure 3). The χ^2 for linear trend (with one degree of freedom) is 27.50, leaving no doubt about its significance. There are, however, worrisome artefactual elements about this trend. First, the lower classes with 3 or 4 extra digits contain a considerable proportion of matings where only one parent was affected, the other being normal. This has been

shown earlier to reduce the incidence of polydactyly. Secondly, the higher the total number of extra digits in the parents, the greater is the proportion of matings, on average, where the female parent is affected on the fore feet, which was shown to increase the incidence of polydactyly.

In an attempt to resolve this question, the effect of number of extra digits on the hind and fore feet separately was examined, excluding all matings where only one parent was affected. The results are summarized also in figure 3. No trend is apparent in either case. This serves to support the contention that the earlier trend, taking all classes of matings into one analysis, was artefactual. However, further analyses require that this statement should be qualified.

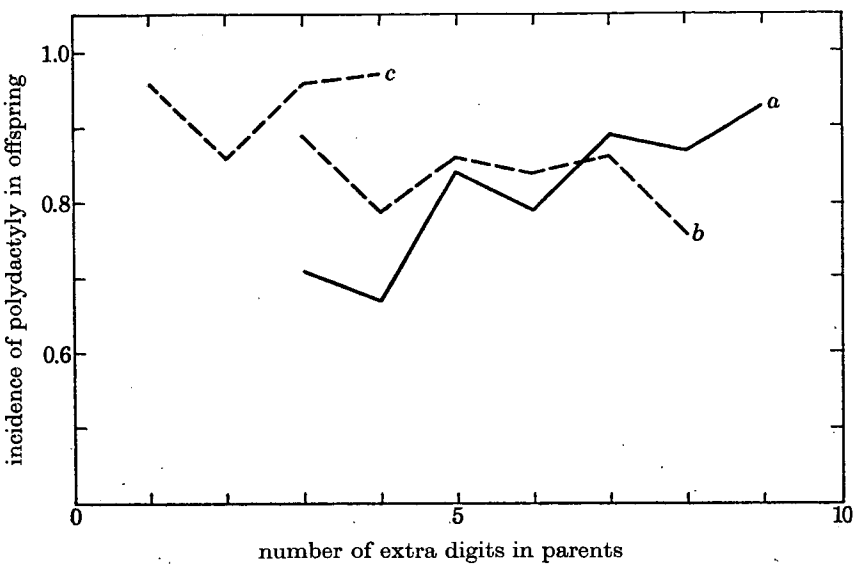


FIGURE 3. Incidence of polydactyly among offspring by varying numbers of supernumerary digits in parents: *a*, total extra digits on all feet of both parents (trend χ^2 for 1 d.f. = 27.50, $P = 0.001$); *b*, total extra digits on hind feet of both parents; *c*, total extra digits on fore feet of female parent.

The untidiness of the system is revealed when we consider separately the matings involving fore-foot polydactyly (in the female parent) and the matings where both parents are affected on the hind feet only. The results are summarized in figure 4, where the appropriate χ^2 values for the trends are shown in the legend. Briefly, among matings where the female parent is polydactylous on the fore foot, there is no connection between the number of extra digits in the parent and the incidence of polydactyly among the progeny (lines *c* and *d* in figure 4). In contrast, where both parents are affected only on the hind feet, the greater the number of extra digits, the higher the incidence (lines *a* and *b* in figure 4).

An explanation of this finding, in terms of the liability model, is that there

are successive thresholds for increasing digit number on the hind feet. This explains the effectiveness of the early part of the selection programme. We could conceptually postulate a similar system for the number of digits on the fore feet, at a higher level than the thresholds for the hind feet, but with the qualification that these thresholds are so tightly bunched that they cannot in practice be distinguished.

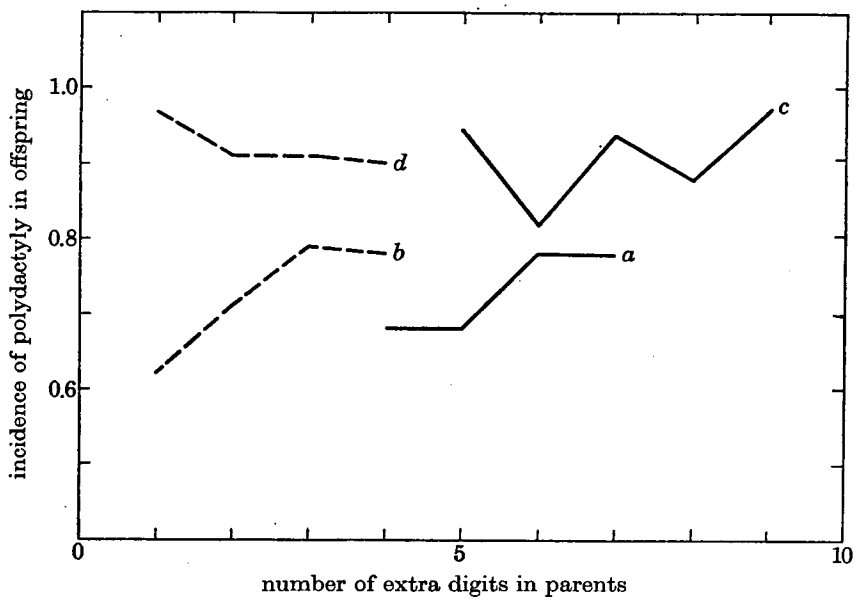


FIGURE 4. Incidence of polydactyly among offspring by varying digit number in parents of different mating types:

	trend χ^2, 1 d.f.
a, female parent unaffected on fore feet, by total number in both parents	4.476
b, female parent unaffected on fore feet, by extra number in male parent only	5.964
c, female parent affected on fore feet, by total number in both parents	0.144
d, female parent affected on fore feet, by extra number in male parent only (slightly negative trend)	1.757

Effect on digit number in offspring. So far, we have discussed merely the incidence of polydactyly among the offspring of parents with different numbers of digits. A more sensitive analysis might be to examine the actual numbers of supernumerary digits in the offspring. The analyses were conducted within the mating types previously isolated to examine trends in incidence. The results are summarized in table 5.

Let us consider first the right-hand side of table 5, i.e. matings involving fore-foot polydactyly. The number of extra digits in the parents was classified in two ways: by the total number on all feet in both parents and by the number on the fore feet of the female parent alone. In neither case is there a hint that

the number of extra digits on the hind feet of the offspring is influenced by the number in the parents. The fore feet of the offspring do not seem to be influenced either, except when the parental number reaches a total of 9, or 4 if we confine our attention to the female parent only. Both of these parental classes include one mating whose 10 offspring had a mean of 3.4 extra digits on the fore feet. This was exceptionally high, the next highest family mean being less than 2. If this mating is excluded, the remaining ones would cause no suspicion of an elevated number of extra digits in these classes. We are aware of the dangers

TABLE 5. THE NUMBERS OF SUPERNUMERARY DIGITS IN THE OFFSPRING OF VARIOUS MATING TYPES

total extra digits in parents	from matings polydactylous on hind feet only			from matings where female parent was polydactylous on fore feet		
	n†	hind feet	fore feet	n†	hind feet	fore feet
4	101	1.584 ± 0.132	0.238 ± 0.070			
5	71	1.662 ± 0.162	0.296 ± 0.095	90	2.389 ± 0.099	0.713 ± 0.110
6	176	1.886 ± 0.094	0.574 ± 0.083	68	1.985 ± 0.144	0.647 ± 0.130
7	128	2.008 ± 0.112	0.414 ± 0.081	195	2.369 ± 0.073	0.687 ± 0.083
8				159	2.195 ± 0.088	0.781 ± 0.095
9				69	2.362 ± 0.101	1.290 ± 0.180

extra digits on the fore feet of ♀ parent	as above, but classified by number on fore feet by parents (i.e. female parent only)		
	n†	hind feet	fore feet
1	121	2.430 ± 0.088	0.901 ± 0.116
2	260	2.093 ± 0.070	0.641 ± 0.068
3	171	2.456 ± 0.074	0.819 ± 0.095
4	29	2.241 ± 0.146	1.379 ± 0.250

† n is the number classified in each category, with the pooling of some extreme categories with small numbers.

of excluding troublesome data, but maintain that on this occasion the deviant observation was truly exceptional. If this is accepted, the balance of evidence is against any influence of parental digit number on the number of extra digits in the offspring, when the female parent is polydactylous on the fore feet.

Turning now to the matings not involving fore-foot polydactyly, on the left of table 5, we meet a contrasting result. The trend in the number of extra digits on the hind feet of the offspring is confirmed by the analysis of variance summarized in table 6. The evidence fails to confirm any real trend on the fore feet (from parents affected on the hind feet only), though there is heterogeneity between groups. In the case of the hind feet, the weighted regression of number of extra digits in the offspring on the midparental values, corresponding to the heritability, is 0.297 ± 0.034. Thus, a heritability of about 30 % for digit number

explains the effectiveness of the early selection based on this criterion, and substantiates the conclusion we reached earlier.

This section has examined whether the number of supernumerary digits can be regarded as an indicator of genetic severity. This is certainly so among mice showing polydactyly on the hind feet only. But once the fore-foot threshold is exceeded, the number of supernumerary digits no longer serves as a measure of genetic severity; the system then behaves as if digit number was mostly developmental 'noise'.

TABLE 6. ANALYSIS OF VARIANCE OF DIGIT NUMBER IN OFFSPRING FROM PARENTS AFFECTED ON THE HIND FEET ONLY

Groups are classified by parental digit number, and the group means are shown in table 5.

		hind feet		fore feet	
source	d.f.	s.s.	m.s.	s.s.	m.s.
total	475	795.8214	—	425.8046	—
within groups	472	783.1414	1.6592	417.1802	0.8839
between groups	3	12.6800	4.2267†	8.6244	2.8748‡
linear trend	1	12.3594	12.3594‡	3.5627	3.5627
residual	2	0.3206	0.1603	5.0617	2.5308

† (Almost) significant at 5 % level.
‡ Significant at around 2 % level.

(iv) *Systemic disturbances of digit number*

The proceeding section left us with an anomaly: since the genetic effect of increasing parental digit number was clearly established when the parents were affected on the hind feet only, why was the effect not seen when one parent (but not the other) was affected on the fore feet as well? Indeed, the trend for the hind feet, in figure 3, is slightly (though insignificantly) negative, even though this group of mice contained those from which a positive trend was established.

The answer derives from the following consideration. The total supernumerary digits on the hind feet (in both parents) could be as high as eight, while the number of the fore feet of the female parent could range up to four. However, when the digits were counted on all feet of both parents, high cumulative values up to the expected 12 were not found. Though 54 fertile matings involving fore-foot polydactyly were recorded, only three matings had a total parental digit number exceeding 9 (two with 10, and one with 11). The reason for this can be seen from table 7, which shows that overwhelmingly, fore-feet polydactyly occurred in those mice which had only one extra digit on each of the two hind feet. 70 % of all recorded cases of fore-foot polydactyly had this particular classification for the hind feet. When the number of extra digits rose to 2 on each hind foot (last row in table 7), the incidence of fore-foot polydactyly fell. Those grossly asymmetrical cases with two extra digits on one hind foot and none

on the other, with the same total number as those with one on each, had a particularly low incidence of fore-foot polydactyly. Thus the distribution of the digits on the hind feet is all-important, while their total number is not. This explains the anomaly posed above; many of the mice with just two extra digits on the hind feet were polydactylous on the fore feet, and therefore severely affected genetically. This was reflected in the high incidence among their offspring, which obscured the underlying trend among mice affected on the hind feet only.

TABLE 7. INCIDENCE OF FORE-FOOT POLYDACTYLY ACCORDING TO THE NUMBER OF EXTRA DIGITS ON THE HIND FEET

hind foot classification†	number of mice	number with fore-foot polydactyly	incidence ± standard error of fore-foot polydactyly
00	160	0	0
01 or 10	71	9	0.127 ± 0.039
02 or 20	20	1	0.050 ± 0.049
11	413	203	0.492 ± 0.025
12 or 21	236	66	0.281 ± 0.029
22	117	11	0.093 ± 0.027
total	1017	290	0

† First number refers to number of extra digits on right hind foot, second number to left foot. There was no heterogeneity among pooled classes.

Although we do not present the details, the tendency towards symmetry of hind-foot manifestation, established earlier (see table 2), did not arise from the inclusion of the (1, 1) class associated with fore-foot polydactyly. The symmetry remained when the data were limited to those mice affected on the hind feet only.

In developmental terms, we are more at ease with the finding if we express it the other way around i.e. that the fore-foot condition affects the number (disregarding for now the distribution) of digits on the hind feet. The data have been reassembled to demonstrate this effect in table 8. There is no doubt about the reduction in the number of extra digits on the hind feet once the fore feet become affected.

We cannot comment profitably on the developmental significance of this negative association between digit number on the fore and hind feet. Certainly, Grüneberg's (1964) suggestion of a competitive system, mentioned earlier, seems helpful. But whether the competition is systemic for 'digit material' or whether it is limited to competition between the hind foot and its associated abnormal hind zeugopodium (a consequence of fore foot polydactyly), we cannot tell. The relative importance of systemic and local influences on limb development remains obscure; the only clear evidence is that both exist. Of more general concern to us here is the demonstration that problems of internal integration influence the degree of manifestation of polydactyly. The observable degrees obviously do

not conform to a neat progressive scale. We have shown one clear case of a disruption of scale, and we therefore cannot dismiss the possible existence of others. The extent to which this vitiates the application of ordinary biometrical techniques is perhaps arguable; what is not in doubt is that it deprives them of any sophistication. This is the reason why we have limited ourselves mostly to descriptive statements.

TABLE 8. NUMBER OF EXTRA DIGITS ON THE HIND FEET, CLASSIFIED BY THE NUMBER OF EXTRA DIGITS ON THE FORE FEET

Data as for table 7, excluding 160 unaffected mice.

no. extra digits on fore feet	no. of mice	no. extra digits on hind feet ± standard error
0	567	2.563 ± 0.037
1	74	2.230 ± 0.078
2	115	2.278 ± 0.047
3	69	2.275 ± 0.068
4	32	2.344 ± 0.115

CONCLUSIONS

The results were to some extent discussed when the data were presented, while their general context was established in the Introduction. This section merely summarizes the main conclusions.

The genetic analysis of the polydactyly is readily compatible with a straightforward model, based on Falconer's (1965) concept of liability, as developed from Wright's (1934) threshold model. We may suppose that, despite the lack of proof, a mutation at a major locus first raised the mean of the population to a level sufficiently close to the threshold to allow the selection to be effective. This initial threshold was for hind-foot polydactyly alone. A higher threshold for a more severe form of the abnormality, involving the fore feet and other skeletal deformities, was also established. While the two-threshold model derived unambiguously from the data, more detailed analyses indicated that it sufficed only as a sketchy outline of the situation. It seems reasonably clear that the hind-foot condition did not correspond to a clean all-or-none division, but that there were multiple thresholds in this region corresponding to increasing numbers of supernumerary digits. The evidence failed to reveal a similar cluster of thresholds for degrees of severity on the fore feet.

In a sense, it is largely irrelevant whether the initial manifestation of polydactyly was the result of a mutation at a major locus or not. Whatever its origin, subsequent selection successfully exploited variation at modifying loci; it was this variation that generated manifold degrees of abnormality and became the main subject of study. The existence of so much latent variation, in a normally

invariant character, leads directly to evolutionary and developmental considerations. These are outwith the scope of this paper, but some of the developmental aspects were mentioned briefly in earlier sections. Some of the genetic variants behaved almost as if they coded for extra digits in specific locations, e.g. on the hind feet, and preferably in equal numbers. More generally, however, the system is more easily described in terms of surplus 'digit material', with some constraints on its distribution between limbs. In particular, the earlier-developing fore limbs, if they became affected, used up more than their fair share. But to attempt a description in these terms is to use words very loosely. The system certainly suggested competitive phenomena, with consequent interactions between systemic and local influences. It is further true that the skeletal disturbance was by no means confined to the feet, and the primary lesion was no doubt far removed from the polydactyly, which was only one aspect of its phenotypic expression. A more complete explanation of the data reported here must await a fuller understanding of the gene control of limb development.

References

Bodmer, W. F. 1960 Interaction of modifiers: the effect of *pallid* and *fidget* on polydactyly in the mouse. *Heredity* **14**, 445–448.

Carter, T. C. 1951 The genetics of luxate mice. I. Morphological abnormalities of heterozygotes and homozygotes. *J. Genet.* **50**, 277–299.

Falconer, D. S. 1965 The inheritance of liability to certain diseases, estimated from the incidence among relatives. *Ann. Hum. Genet.* **29**, 51–76.

Fisher, R. A. 1950 Polydactyly in mice. *Nature, Lond.* **165**, 407.

Forsthoefel, P. F. 1962 Genetics and manifold effects of Strong's luxoid gene in the mouse, including its interactions with Green's luxoid and Carter's luxate genes. *J. Morph.* **110**, 391–420.

Grüneberg, G. 1964 The genetics of skeletal abnormalities. Proc. 2nd Int. Conf. on Congenital Malformations (New York, 14–19 July 1963). Edited and published by International Medical Congress, Ltd.

Holt, S. B. 1945 A polydactyl gene in mice capable of nearly regular manifestation. *Ann. Eugen.* **12**, 220–249.

Johnson, D. R. 1969*a* Brachyphalangy, an allele of extra-toes in the mouse. *Genet. Res.* **13**, 275–280.

Johnson, D. R. 1969*b* Polysyndactyly, a new mutant gene in the mouse. *J. Embryol. exp. Morph.* **21**, 285–294.

Roberts, R. C. 1967 The limits to artificial selection for body weight in the mouse. IV. Sources of new genetic variance – irradiation and outcrossing. *Genet. Res.* **9**, 87–98.

Wright, S. 1934 The results of crosses between inbred strains of guinea-pigs, differing in number of digits. *Genetics* **19**, 537–551.

PAPER 22

Small eyes - a new dominant eye mutant in the mouse.

Genet. Res. Camb. 9, 121-122. 1967

by

R.C. ROBERTS.

Genet. Res., Camb. (1967), 9, *pp.* 121–122
Printed in Great Britain

Small eyes—a new dominant eye mutant in the mouse

By R. C. ROBERTS

A.R.C. Unit of Animal Genetics,
Institute of Animal Genetics, Edinburgh, 9

(*Received* 18 *August* 1966)

In the latest issue of *Mouse News Letter* (No. 35, July, 1966), a comprehensive listing of mutant gene symbols classifies twenty different mutants as being most readily recognized by their effects on the eye. A new eye mutant has been found in this laboratory, and the first problem is to decide whether it is different from the ones already known. The new mutant is dominant, which conveniently distinguishes it from sixteen of the known twenty. Judging from the descriptions of the effects of the remaining four, which are dominant, it is also clearly different from three of them. The new mutant described here bears several superficial resemblances to *Blind*, described by Vankin (1956). However, *Blind* has its eyelids open at birth, whereas the one reported in this note has its eyelids closed until the normal age. It thus appears to be a mutant not previously reported.

Since the new mutant has no obvious effect on the mouse other than to reduce the size of the eye, it has been called *Small eyes*. This distinguishes it from other named mutants in the mouse, and by the same criterion, the symbol *Sey* is proposed for it.

The mutant first arose in a line of mice previously selected for low body weight in which selection had been suspended for nine generations. In one mating, a first litter of seven mice contained two females and one male whose eyes were obviously smaller than normal. No further litters were obtained from this mating; both parents had normal-looking eyes. The affected progeny were mated together and to their normal sibs. When these matings generated suggestive segregation ratios, the stock was expanded and further test matings were made, with results as shown in Table 1.

Table 1

Phenotype of		No. of offspring classified	Phenotype of offspring	
♀ parent	♂ parent		*Sey*	+
Sey	*Sey*	154	96	58
+	*Sey*	33	18	15
(a) *Sey*	+	74	34	40
(b) *Sey*	+	32	14	18

(a) unaffected males from same stock
(b) unaffected males from unrelated stock

The following conclusions can be made from the data:

1. The mutant is dominant; this is confirmed in the table from the matings of *Sey* ♀♀ by males from an unrelated stock.

2. The mutant is lethal in its homozygous state. Though the embryology has not so far been examined, no other postulate will explain the following combination of facts:

(*a*) *Sey* × *Sey* gives a segregation ratio that differs from expectation based on a 3:1 ratio at the 0·1% level. But it shows excellent agreement with a 2:1 ratio.

(*b*) No matings have yet been found which give all *Sey* progeny.

(*c*) Pooled matings of *Sey* by normal give a segregation ratio of 1:1.

3. Agreement with the expected ratios is such that it arouses no suspicion concerning the viability of *Sey* heterozygotes.

The expressivity of the mutant phenotype is variable. Often, one eye is more seriously affected than the other. Sometimes, there is but little reduction in the size of the eye compared with the normal; at the other extreme, the eye is very small and remains tightly closed. But from preliminary investigations, it never quite reaches a stage that could be described as anophthalmia.

Matings between unaffected animals from the stock gave all normal offspring, as expected from a dominant gene, with one troublesome exception. This mating was between supposedly unaffected animals, but produced twelve *Sey* progeny and seventeen normal ones. When the parents were re-examined later, it was easy to imagine that the female parent had at least one of its eyes smaller than normal. However, the technician in charge of the stock disagreed, and reclassified this female as normal when she was mixed with other mice of similar phenotypes, from the same stock, the mixed lot containing normal and slightly affected animals. There is therefore a slight doubt about the full penetrance of mutant. But errors of classification on this account cannot be serious, otherwise the segregation ratios in all of the backcross matings would not have fitted so well.

Sey has not been tested systematically for linkage. It was, however, tested against *microphthalmia*, (*mi*), using the dominant allele Mi^{wh}. The recombination frequency was $0{\cdot}55 \pm 0{\cdot}04$.

This short note describes the preliminary findings. The mutant has now been taken over by Mrs R. M. Clayton and her collaborators for a more detailed investigation of its effects.

SUMMARY

A new mutant, *Small eyes* (symbol *Sey*), that reduces the size of the eye in the mouse, is described. It is dominant, and lethal when homozygous.

I am indebted to Miss Carol Sergeant, who found the mutant in a stock of mice under her care.

REFERENCE

VANKIN, L. (1956). The embryonic effects of 'Blind', a new early lethal mutation in mice. *Anat. Rec.* **125**, 648.

PAPER 23.

Some contributions of the laboratory mouse to animal breeding research.

Part I: Growth

Animal Breeding Abstracts, 33 (3), 339-353. 1965.

by

R.C. ROBERTS

Commonwealth Bureau of Animal Breeding and Genetics

REPRINT No. 66

Some Contributions of the Laboratory Mouse to Animal Breeding Research

BY

R. C. ROBERTS

A.R.C. Unit of Animal Genetics, Institute of Animal Genetics, Edinburgh, 9

Reprinted from

ANIMAL BREEDING ABSTRACTS, Vol. 33, Nos. 3-4, pp. 339-353, 515-526, September-December 1965

SOME CONTRIBUTIONS OF THE LABORATORY MOUSE TO ANIMAL BREEDING RESEARCH

PART I

R. C. ROBERTS
A.R.C. Unit of Animal Genetics, Institute of Animal Genetics, Edinburgh 9

The current interest in the field of molecular biology, where experimental evidence from micro-organisms is frequently deployed to elucidate the nature of the gene, serves as a reminder that the science of genetics transcends species barriers, and emphasises the ubiquity of basic genetical principles. In this context, it is perhaps unnecessary to justify the examination of laboratory animal research for its relevance to the breeding of domestic livestock. Indeed, this relevance is now widely accepted, and no animal production conference today seems complete without a session devoted to laboratory animals. It therefore seems pertinent that a review should be attempted of some salient features of research with the laboratory mouse, in so far as it may affect the thinking and activities of research workers whose concern is the breeding of farm animals. The volume of material available makes it necessary to restrict the review to questions of genetic analysis, and the vast literature on reproductive physiology, for instance, will be left untouched.

We may pause briefly to consider what kind of information from mouse work may be of interest to the breeder of livestock. Basically, many questions concerning animal breeding are common to all species. Stated concisely, the main requirement in practice is to describe the " genetic architecture " of various characters. This means a knowledge of the extent to which genes affect the phenotypic expression of a character—how its total variance is partitioned into various genetic and environmental components. Given this, the modern theory of quantitative inheritance is sufficiently well established to indicate how best the genes may be manipulated to change the expression of the character in any desired direction, or indeed whether it may be changed at all. Thus we may decide whether selection is likely to be effective, and which form of selection is likely to have the greatest efficacy. Or we may decide that inbreeding, followed by the crossing of inbred lines, may be a more efficient method. A knowledge of the genetic correlations involved enables us to predict, to some extent, how a change in one character may be expected to bring about concomitant changes in others. However, the most profitable utilisation of this approach demands an accurate knowledge of the relevant genetic parameters for the species concerned. In this respect, the laboratory mouse may not always be a particularly good analogue for farm animals, and the application of detailed results across species should obviously be exercised with caution. In this sense, mouse results should be regarded as a provisional indication of a possible outcome. It is in the nature of scientific progress that generalities become established only through time, as the volume of knowledge increases.

For a mouse experiment to be useful to the breeder of farm animals, it must aim to elucidate the genetic control of a particular situation, rather than devote itself to the operational pursuance of some quantitative measurement. If the objective is, say, to decrease the backfat measurement in the pig, then there is little point in doing the work on any organism other than the pig, and very little genetics is involved. But if the objective is to discover the genetic control of various fat deposits in the pig, then any relevant or analogous work on the mouse would be of obvious interest. Mouse research may find a particular application to animal breeding research as a pilot study in an unknown situation. It may find a more general application in the establishment of the kinds of genetic properties to be expected for different kinds of characters. If, as an unlikely instance, brain cholinesterase levels in hill sheep were to be shown to affect their grazing behaviour, the considerable knowledge of the inheritance of cholinesterase level in laboratory animals would immediately have some implications for sheep breeders.

As examples of the specific kinds of information from mouse populations that may be of interest to the breeders of livestock, we may cite the responses to selection—their pattern, duration and

ultimate cessation. If the limit has been reached, what is its nature, in genetic terms, and how may it be overcome, if at all? How is the response to selection modified by the mating system employed, or the environment in which it is practised? What level of inbreeding is necessary before crosses between lines can profitably be exploited? Is specific combining ability likely to be important, or can we predict, with reasonable accuracy, cross performance from pure line performance? Experiments with laboratory animals ought to be able to offer some guidance to the livestock breeder on these and similar questions. Not all of them have yet been fully answered.

The employment of the laboratory mouse for some of these studies needs little justification. Though the applicability of mouse results to the larger animals must obviously be checked before complete confidence can be placed in that application, preliminary results may be obtained with a relatively small expenditure and in a relatively short time. Indeed, for problems like the limits to selection, we are forced to use laboratory animals; with the possible exception of some poultry stocks, it is probable that no breed of farm animal has yet been selected to the limit for any trait. Of course, if the concern is merely with cost and time, then the mouse cannot compete with the fruit fly, *Drosophila*, on which so many experimental studies of quantitative variation have been done. Mouse workers also lack the sophisticated techniques available to *Drosophila* workers, whereby chromosomes with a specified gene array can be synthesised almost at will. The advantages of the mouse over *Drosophila* stem from its biological organisation, which is much more akin to that of farm animals. In particular, the mouse gestates and suckles its young, and thus has a series of traits, like litter size and lactation, whose analogues are of obvious economic importance in domestic livestock. To the extent that a mouse is more similar to a cow or a pig than is a *Drosophila*, results from the mouse may be that much more readily applied to the larger animals.

Against this background, we shall examine some of the experimental work with the mouse on characters that bear at least a superficial resemblance to characters of economic importance in domestic livestock. The reader is referred also to the reviews of Chapman (1951, 1961), which cover some of the earlier work with laboratory animals.

I. Genetic Analysis of Growth and Body Weight

Selection for body weight

Live weight gain and weight for age are among the most important of economic traits in farm animals, and this may, in part, have been the motivation for the considerable volume of work on the weight of the mouse. But body weight is also easy to measure, which is a strong pragmatic reason for choosing it as a character in experimental studies.

There is by now ample evidence of great variation in the weight of the mouse at a given age. Judging from a study by Crowcroft and Rowe (1961), this is true even of wild mice, when captured animals are supplied with excess food in captivity. A demonstration of the additive genetic nature of at least part of the variation in the laboratory mouse was supplied by Goodale as early as 1938. Using a form of progeny testing, he selected mice for large body weight, and reported progress over 14 " chronological groups ", which corresponded roughly to generations. A second report, which Goodale published in 1941, seemed to indicate that this response had continued more or less linearly over a further 14 generations, making a total of 28. Though Goodale had no controls, his males by this time averaged about 43 g. at 60 days of age, whereas his starting point was about 25 g. Goodale speculated that the potential power of selective breeding was enormous, and he clearly expected further progress. However, Falconer and King (1953) examined his data and some of his later results, and concluded that, by the time of his 1941 report, he had in fact more or less attained his maximum response. In retrospect, the main feature of Goodale's experiment is that it was the first to establish the feasibility of selecting for body weight in the mouse.

Not long after Goodale, MacArthur (1944*a*, 1949) reported a selection experiment for both high and low 60-day weight. The 1949 paper showed the results for 23 generations, by which time the large mice (about 38 g.) were roughly three times the size of the small ones (about 12 g.). By then, MacArthur found evidence that his response was already tailing off, and Butler (1952), working with the same strains, found that, in fact, little if any further progress was made. A portentous feature of MacArthur's results was the correlated responses to his selection for weight, which were apparent after even eight generations of selection. He reported on these in some detail (MacArthur, 1944*b*; MacArthur and Chiasson, 1945). Briefly, he found that the large line was more docile and inactive than the small, had comparatively shorter ears, feet and tail, and had a higher ovulation

rate. Colour gene differences had also appeared between the lines, but MacArthur himself seemed happy to accept random genetic drift as the probable explanation here.

As a footnote to MacArthur's work, we can note with satisfaction his view of the genetic basis of selection. Though he had adopted a cumbersome combination of mass and sib selection, supplemented by some informal progeny testing, he was perfectly clear about his objective. In his words—" The job was to re-shuffle the size genes ". Others may have defined the aims of selection equally clearly, but no one has put it more succinctly. Goodale too, in his 1938 paper, had meditated on the task of amassing all the potentially useful genes in one stock, and though he failed to abide by his own dictum, he emphasised the need to select from a large initial population.

MacArthur's strains were employed in further studies designed to elucidate the genetic control of body weight. Butler (1952) crossed the large and small strains together, and to two other inbred strains, and analysed the crosses. In all cases, he found the F_1 and F_2 means to be intermediate between the parents, while the backcrosses were intermediate between the F_1 and the respective parental strain. Thus, the mean weight showed proportionality to the percentage of genes for large size in the genome. The F_2 variance was no greater than that of the F_1. Though Butler at the time expressed surprise at this finding, it can be shown (*see* Falconer, 1953) that the increase in variance to be expected is of the order of $1 \cdot 6\ g^2$; this increase might not be noticed, as it is a relatively small proportion of the total variance.

Results essentially similar to Butler's were reported by Chai (1956*b*), though Chai found some increase in the variance of the F_2 over the F_1 level.

A second study involving MacArthur's strains was that of Lewis and Warwick (1953). They crossed both the large and the small strains to an unselected randombred strain from the same base as the selected strains. They then backcrossed to the large strain and continued selection for high 60-day weight; similarly, a backcross to the small strain was selected for low weight. In each case, they selected with both outbreeding and inbreeding mating systems. They observed responses over a further five generations in both directions, and the results were essentially similar in both inbred and outbred populations. In view of Butler's (1952) finding that MacArthur's original strains had by this time ceased to respond, we must adduce from Lewis and Warwick's results that an infusion of genes from the base population had been responsible for this renewed response. The same authors (Warwick and Lewis, 1954*a*) crossed their large and small strains; their results agreed with those of Butler, quoted previously.

The quantitative nature of the genetic variation in body weight of the mouse was thus firmly established from some of the early work. However, we may cite here an example of the disturbance that may be occasioned to the distribution of a quantitative trait by a single gene substitution. MacArthur (1944*a*) found animals that he described as " runts " in his small strain. Lewis and Warwick (1953) also found them in the small strain, and noted that they greatly increased the variance. However, King (1950, 1955), who was also working with MacArthur's small strain, showed that these " runts " resulted from the action of a single recessive gene, which King labelled *pygmy*. Homozygous *pygmies* at six weeks old are approximately half the size of normal litter mates, irrespective of the genetic background for size. Furthermore, *pygmies* are either completely sterile or have a much reduced fertility. But it appears that the gene may not be completely recessive with respect to size. Warwick and Lewis (1954*b*), in a special study, concluded that heterozygous individuals were, on the average, sufficiently smaller than homozygous normals, to have had approximately twice the chance of being selected for breeding purposes in the small line, under the conditions of their experiment. This indicates how a deleterious gene may have a selective advantage in a population by virtue of a slight effect in the heterozygous state.

The next study of the effect of selection on body weight was that reported by Falconer (1953), whose genetic analysis of the situation was more sophisticated than any published previously. Falconer selected for high and low six-week weight, thus permitting a shorter generation interval than that achieved by the earlier workers who used 60-day weight as their criterion of size. His base population was a four-way cross of inbred strains; in terms of genetic variation, this cross is the equivalent of a single full-sib family from a randombred population. Thus the population was not broad-based, in genetic terms. Falconer's method of selection was radically different from that practised by Goodale or MacArthur, in that animals were selected on the basis of their deviation from the litter mean for their own sex. Each family, as far as possible, contributed one male and one female as parents in the succeeding generation. This within-family method of selection was an important development in laboratory work, for though it utilises only half of the additive genetic variance, this restriction is outweighed by two important advantages. Firstly, it circumvents any

variation due to maternal effects and thus simplifies the genetic interpretation of the data. Secondly, by doubling the effective population size it greatly reduces the rate of inbreeding, and thus allows experiments to be conducted with less space and labour.

Falconer found that his high and low lines diverged regularly over the eleven generations which he first reported. However, this response was markedly asymmetrical. Whereas his high line gained a total of 4 g. in mean weight, the low line decreased in weight by 7 g.—nearly twice as much. This asymmetry was reflected in the heritability estimates of 0·20 for upward selection and 0·50 for downward. A more detailed analysis of the asymmetry, with further results, was given in a later paper (Falconer, 1955). Six-week weight can be regarded as consisting of two parts—the weaning weight at three weeks, which is largely a characteristic of the mother, and the growth from three to six weeks. When Falconer examined these two components separately, he found that the asymmetry of response was entirely attributable to weaning weight and not at all to post-weaning growth. As a possible explanation of the asymmetry in weaning weight, Falconer hypothesised that the maternal component itself consisted of two parts—one being directly related to maternal body size, while the other is independent of body weight at the time but adversely affected by a deviation of weight, in either direction, from the original level. Thus these two effects would cancel each other in the high line, but act in conjunction in the low so as to reduce weight.

In his earlier paper, Falconer argued convincingly that the asymmetry could be explained largely by the rise in the inbreeding coefficient during selection, which had reached 38% by the eleventh generation. As he found his stocks to exhibit directional dominance in favour of large size, inbreeding depression would operate adversely in the high line, while in the low line it would assist the selection for small size. Taylor (1954), working with Falconer's stocks, found that this directional dominance was similarly a feature of the three-week weight, which reflects maternal performance. Falconer could therefore have ascribed his later results also to the differential effects of inbreeding, but he seems to have abandoned this possible explanation for no compelling reason. However, in his book, Falconer (1960*a*) enumerates the manifold causes that may result in asymmetrical responses; it is frequently difficult to differentiate between these causes.

Falconer's main conclusions, the asymmetrical response apart, may be summarised as follows:

(i) A large number of loci of approximately equal effects control the genetic variation in body weight.

(ii) Dominance is predominantly in the direction of large size.

(iii) The response ceased in both lines after about 20 generations.

(iv) The large and small lines (about 28 and 12 g., respectively) eventually diverged by sixteen times the original genetic standard deviation.

(v) The realised heritability remained unchanged in both lines until the limit was reached.

(vi) Realised selection differentials equalled the expected differentials in the high line, but fell short in the low, indicating that natural selection impedes progress when selecting for small size.

(vii) When selection was suspended, the small line reverted towards the original level, corroborating the finding with respect to natural selection; the large line did not revert.

(viii) Compared with the small line, the large line had longer tails, had higher twelve-day and three-week weights, and had a larger litter size; the number of fertile matings and post-natal viability fell in both lines.

A recent report by Rahnefeld, Boylan, Comstock and Singh (1963) supports many of these conclusions. Their experimental procedures differed in several important respects from Falconer's. Firstly, their base population was derived from a reciprocal cross of only two inbred strains. Secondly, they selected for growth between three and six weeks. Thirdly, they conducted a mass selection procedure based entirely on the growth of the individual mouse. Selection for increased growth for 17 generations changed the mean growth by 4 or 5 g., depending on how the increase was estimated. This progress, however, represented only about six times the original additive genetic standard deviation. Various heritability estimates ranged from 22 to 26%. Their response was linear over the period of study, and, at the time of reporting, showed no indication of diminishing. The authors conclude that many genes were involved in this response. They also report data showing that dominance is directional towards increased growth. Finally, as growth was increased by selection, a correlated response was observed in favour of an increase in litter size.

It is thus clear that several experiments, employing a variety of selection procedures, have established that much of the variation in the body weight of the mouse has an additive genetic source.

The limits to selection for weight

The cumulative conclusion from all these studies is that two-way selection, over an interval of 20 to 25 generations, brings about a divergence of body weight in laboratory populations of the mouse, representing, perhaps, a factor of eight or ten times the original phenotypic standard deviation. After that, the response may be expected to cease. It is thus pertinent to enquire about the genetic nature of these limits to selection, and to investigate the possibility of further progress. Indeed, Goodale had set up his pioneering experiment in 1930 explicitly " to determine limits of change which can be made by selection " (Goodale, 1938). However, his optimistic expectation of " an indefinite or even an unlimited amount of change " was not substantiated.

An experimental study of the problem was published by Falconer and King (1953). They obtained samples of both Goodale's and MacArthur's original large strains. Continued selection for high body weight confirmed that the strains had ceased to respond, though there was some evidence of a slight response when these large strains were selected downward. Falconer and King noted that whereas Goodale's mice were large-bodied but not very fat, MacArthur's mice were smaller in linear dimensions but were very fat. From this, Falconer and King argued that a cross between the two strains should provide new genetic variance upon which continued selection could act. This expectation was realised in practice, and the mean six-week weight rose from about 29 g., in each of the parental strains, to 32 g. after seven to nine generations of selection from the cross. Downward selection from the crossbred was even more successful—a result reminiscent of Falconer's earlier finding. This downward selection was more successful also than that from either of the two parental strains, confirming that new genetic variance had been generated. Falconer and King's conclusion was that the lack of response in the parent strains could be attributed to the loss of additive genetic variance, and that furthermore, this loss had to be ascribed largely to the fixation of loci by selection. The inbreeding accumulated during that selection could account for only a part of the loss of variance. What their results established beyond doubt was that the two strains were sufficiently differentiated genetically to provide a useful amount of new genetic variance on crossing.

Though it does not fit easily under the rubric of " limits to selection ", it is convenient in this context to refer to another aspect of the genetic effects of selection, as discussed by Falconer and Robertson (1956). It seemed possible that genes affecting the mean expression of a selected character may do so either directly, or by increasing the susceptibility of the individual to environmental sources of variation. Genes of the latter kind would be expected to increase the phenotypic variance of a character selected in either direction, unless they were completely obliterated by genes affecting the character directly. As both MacArthur and Falconer in their studies had noted an increase in the variance of their large lines (though they both noted a decrease in their small lines), it seemed apposite to screen a population experimentally for genes that affected the sensitivity to environmental sources of variation. Falconer and Robertson (1956) conducted this test, using body weight in the mouse as their character, by mating together animals of opposing extreme weights in one line, and animals of intermediate weights in another. It was hoped that thereby the mean weights would neither change nor differ in the two lines; but genes affecting environmental susceptibility would cause an increase in the variance of the " extreme " line compared to the " central " one. The result was completely negative, in that the coefficients of variation of body weight were identical in both lines over 13 generations. Thus, selection of extreme phenotypic deviants did not materially change the sensitivity to environmental influences.

Genotype-environment interactions involving weight

Breeding experiments with the laboratory mouse normally fail to encounter one important feature of livestock breeding. Under laboratory conditions, mouse populations are usually kept in a controlled environment of considerable stability, whereas commercial livestock are subjected to a great variety of environmental conditions. In genetic terms, any genotype-environment interaction may wreck the achievements of selective breeding if an improved breed or strain is transferred to an environment unlike the one where the original selection was made. It therefore became clear that an experimental investigation of these possible interactions was required.

The theoretical groundwork for such experimentation was laid by Falconer (1952), who regarded a phenotypic measurement in two environments as two distinct characters. To the extent that genes favouring a certain phenotype in one environment would also operate favourably in a second environment, the two " characters " would be genetically correlated. Falconer thus framed the question of genotype-environment interaction in terms of a genetic correlation. This permitted the theory to be developed sufficiently to predict the progress expected of a character, selected in one environment, when transferred to another. The theoretical conclusion was that animals should usually be selected for improvement in the environment in which they are destined to live. This conclusion was an important development, for, at the time, ideas of animal breeding were somewhat tinged by Hammond's (1947) opinions that animals should be selected under optimal environmental conditions, so that the fullest expression of the character is favoured. Hammond supposed, further, that animals so selected would retain their superiority when transferred to poorer environments. Falconer and Latyszewski (1952*a*) commented on the mutual contradiction of Hammond's two premises; animals that fail to reveal their superior genotype in a poor environment for the purpose of selection will equally fail to reveal it, in that environment, for the purpose of production, should that genotype be uncovered under better conditions. Hammond's ideas, wittingly or unwittingly, coincided with some of the mythology of pedigree breeding, geared to a show-ring mentality. Hammond's and Falconer's views were obviously in direct conflict, but one amenable to experimental resolution.

Falconer and Latyszewski (1952*b*) reported on one such experiment. Two strains from the same base population were selected for high six-week weight; one strain was fed *ad libitum*, while the other was restricted to about 75% of the normal intake, from weaning at three weeks until six weeks of age. The effect of the restriction of food intake was to reduce the six-week weight by some 10%. However, the weights of both strains increased over eight generations of selection. The crucial tests were performed after the fifth, seventh and eighth generations, when animals selected on one plane of nutrition were measured on the other. The results were similar on each occasion. On the restricted diet, animals selected on that diet were much superior to those selected on the full diet, which failed to show any improvement over the unselected diet. On the full diet, the animals selected on that diet were now superior, but those selected on a restricted food intake were but little lower. Contrary to Hammond's thesis, therefore, this experiment indicated that animals should be selected in the environment where they are to be kept, and to this extent, Falconer's theoretical conclusion was vindicated. But the agreement with theory was not complete, as Falconer's derivations had indicated a symmetrical correlated response. The full-diet strain should therefore have shown considerable improvement when reared on a restricted diet, but it did not. In an attempt to clarify the situation, Falconer (1960*b*) published a second study, which was a more adequate experiment than the first. Again, mice were selected on two planes of nutrition, on this occasion for both high and low growth, between three and six weeks. This time, the restriction of food intake was obtained by diluting the normal diet with indigestible fibre; its effect was to reduce weight by 20%. The diets were exchanged for samples of the two high lines in each generation, with a similar exchange for the two low lines. In all their essential features, the results from the second experiment confirmed those obtained from the first. Responses were observed over 13 generations of selection for both high and low growth on both planes of nutrition. However, when the diets were exchanged, the anomalies persisted. When tested on the full diet, the large mice selected on either plane performed equally well. But on the diluted diet, the large mice selected on the other diet were much inferior. The small mice produced a " mirror-image " effect. On the full diet, the small strain selected on the full diet showed the greater progress; no difference was observed when the two small strains were tested on the low plane. The results therefore confirm that the safest measure is to select in the environment where the animals are required to live. However, should the selection be carried out in one environment, then, if these results have any generality, that selection should be conducted in the environment *least* favourable to the desired expression of the character. Another result of potential importance, observed in both experiments, was that mice selected on the high plane were much fatter, though no heavier, than those selected on the low plane, when both were tested on full diet.

The asymmetrical correlated responses therefore remained unexplained. Changes in the genetic parameters during selection pointed to a complex genetical situation. Recently, however, Bohren, Hill and Robertson (1965) have brought new theoretical considerations to bear on the problem, showing that the correlated responses need not be expected to be symmetrical.

Korkman (1961) reported an experiment similar in design to that of Falconer and Latyszewski's

(1952*b*) study. Korkman's low plane was very low; it was only just adequate to allow the animals to survive. Though Korkman accumulated a considerable selection differential on the low plane, he observed no response whatsoever to selection for high 40-day weight. But on the high plane, he obtained a normal response. The lack of response on the low plane, whatever the cause, makes it difficult to compare Korkman's results with those of Falconer and Latyszewski when the diets were exchanged. Briefly, what Korkman found was that each strain was better than the other on the diet appropriate to its selection. Thus Korkman concluded that performance was best improved by selection on that plane of nutrition on which the performance was to be measured. To this extent, the two studies were in qualitative agreement.

Dalton and Bywater (1963) reported an experiment of a similar type. They selected for the total weight of the litter at weaning time on two diets. However, 14 generations of selection produced no response on either diet, and, not surprisingly, no correlated response was observed when the diets were switched.

Korkman (1957) had reported another interaction experiment of considerable interest involving body weight. In this experiment, the sex of an animal could be regarded as a part of the " environment " in which that animal lived. The question was whether the difference in body weight due to sex exhibited any hereditary variation. To this end, he selected two lines, to maximise and minimise respectively the sex difference in body weight at 90 days of age. In one line, he selected the heaviest male and the lightest female in each litter as the parents of the next generation. In the other line, he selected the lightest male and the heaviest female. Taking the average of the last three of his ten generations of selection, he increased the sex difference from 2·2 to 4·3 g. in the first case, while the difference remained fairly stable in the second. Korkman thus showed that a genotype which promotes growth in the one sex may promote it to a lesser extent in the other.

The studies reported above were all directed towards the " building in " of a genotype-environment interaction through selection in specified environments by producing strains especially adapted to those environments. But other investigators have screened existing genotypes for similar interactions. The first such study to involve weight in the mouse appears to be that reported by Young (1953). Young tested three inbred strains on two diets and in two temperatures. The results were largely negative, though there was some suggestion that one of the three strains, C57BL, had its weight particularly depressed by a diet consisting entirely of crushed oats.

Bakels (1963) developed a different approach to the same problem. Four sire progeny groups, each comprising 334 daughters, were divided evenly between a standard 20% crude protein diet and a diet with 14% protein of vegetable origin. He measured litter weight at 15 days. His analysis of variance shows no significant interaction between sire groups and diet, though Bakels could point to the fact that the number and size of the significant differences between the sire groups were not the same on the two diets. This apparent enigma may be due to the fact that the analysis of variance is not a particularly sensitive technique for this purpose.

Bakels' results perhaps epitomise much of the work on genotype-environment interaction, particularly with respect to body weight in the mouse. Evidence of interaction can usually be produced by a thorough search, but the fact that the search has to be thorough in the first place indicates that, in practice, these interactions are probably not of overriding importance. Alternatively, in order to demonstrate an interaction, we have to look for investigations where the difference between the environments is extreme. Barnett (1965) has recently summarised his extensive publications on the adaptation of mice to a cold environment. He has shown that inbred strains adapt differentially to a cold environment, where the temperature is —3° C compared to a normal 21° C. As one example, his A2G strain after many generations in the cold had a body weight equal to the control level, though the effect of the cold initially was to reduce body weight. A peculiar feature of this adaptation was its accumulation over many generations. Barnett stresses that heterozygosity confers much resistance on animals to the effects of a transfer to a cold environment.

Complementary to some of Barnett's results, we may note that Harrison (1963) reports that two inbred strains of mice reacted differently, with respect to body weight, when transferred to a hot environment (32° C).

As a final example of an interaction involving diet, though it does not bear directly on weight, we may take the experiment reported by McNutt and Dill (1963). This study illustrates how a drastic modification of the diet may well generate pronounced interactions. McNutt and Dill provided mice with 4% sodium chloride solution as their sole supply of drinking water. Whereas one strain (NH) reached 50% mortality after one week under this regime, a second strain (IHB)

did not reach the same level of mortality until after 35 weeks of treatment, and could therefore tolerate a high salt intake over a prolonged period. In a genetic investigation, McNutt and Dill found that resistance was dominant over sensitivity to the effect of high salt intake, and that the difference between the two strains in this respect was controlled by relatively few loci.

The main conclusion, then, with respect to genotype-environment interactions as measured by body weight, is that these interactions do not seem to be important, unless a severe modification of the environment is invented. But before possible interactions are too lightly dismissed, one general point deserves to be made. Young, in his 1953 paper, pointed out that a vast array of both genotypes and environments is possible, and that a single study can therefore include only a small fragment of these. Any one experiment on interactions hence lacks generality *ab initio*, and does not necessarily preclude interactions if either different genotypes or different environments are sampled for further studies. With this proviso, the main conclusion stands.

Correlated responses to selection for weight

Selection for body weight in the mouse has invariably produced concomitant changes in other characters. Indications of this have already been quoted. Some of these correlated responses have been studied in greater detail.

The primary implication of a correlated response is that the character under selection is genetically correlated with another character, though other parameters affect the magnitude of the correlated response. If the relevant parameters are known, then the expected correlated change can be predicted from theory. The first experimental verification of the theory, from an experiment with mice, was reported by Falconer (1954). Previous studies had shown that large animals had longer tails than animals selected for small size. Falconer therefore selected one pair of lines for large and small size (using six-week weight as his criterion), while another pair of lines from the same base population was selected for long and short tails, respectively. For each pair of lines, he observed both the direct response and the correlated response in the other character. This gave him two estimates of the genetic correlation (0·62 and 0·57), which were in excellent agreement. From this, Falconer concluded that the theoretical treatment adequately accounted for the observed correlated responses to selection in his experiment.

Cockrem (1959) used the same two characters in another population of mice to determine whether, despite the positive genetic correlation, selection could be applied to increase one trait while decreasing the other. He estimated the " expected body weight " from a regression equation of weight on tail length. He then selected for both positive and negative deviations from the regression line, *i.e.* he selected for animals with greater body weight and shorter tails in one line, and smaller body weight with longer tails in the other. Cockrem found a fairly rapid response over six generations of selection. For instance, the males in this line selected for positive deviations ended up with body weights on $3\frac{1}{2}$ g. greater than the other line, while their tails were some 2 cm. shorter. Thus Cockrem established that the presence of a genetic correlation of a fairly high order does not preclude the possibility of selection for various combinations of the correlated traits.

By way of incidental comparison with the results of Barnett and of Harrison (quoted earlier) on interactions between genotype and temperature levels, Cockrem's (1963) study is of interest. He kept his two strains, with their relatively longer and shorter tails, at three temperatures, 7°, 21° and 32° C, from three to six weeks of age. He found no evidence of any interaction of strain by temperature level, with respect to body weight. But the line with the longer tails showed a relatively greater increase of tail length under the hot environment.

It was noted earlier that increased size in the mouse is accompanied by an increase in the ovulation rate, which is reflected in a bigger litter size. A closer examination of the factors involved was described by Fowler and Edwards (1960). They obtained data from two unrelated sets of large and small strains—the N strains, described by Falconer (1953), and the C strains, described also by Falconer (1960*b*). Fowler and Edwards confirmed that the large mice had a higher ovulation rate, roughly twice that of the small mice. The relationship between body weight and ovulation rate was linear, and identical in both sets of strains after body weight had been corrected for the weight of carcass fat. But this linear regression, though obvious enough when all the data were taken together, did not necessarily hold within strains. Fowler and Edwards found considerable sterility in the N strains, both the large and the small, though the fertility of the C strains had been unimpaired by selection. In the large N strain, sterility was attributed to low libido of the males; sterility in the small N strain was due to a hypo-functioning of the anterior pituitary in some females.

An interesting aspect of Fowler and Edwards' results was the differential responses of the large and small strains to superovulation techniques. This indicated a difference between large and small mice in the amounts of endogenous pituitary gonadotropins secreted. Since the pituitary gland is profoundly involved in the physiology of growth, selection responses for growth may well be mediated by changing the endocrine activity of the pituitary. This possibility was examined by Edwards (1962). He found a close correlation between body weight and the weight of the pituitary, but the weight of pituitary per unit of body weight was identical in both large and small mice. Edwards next examined the endocrine activity per unit weight of pituitary tissue by a bioassay technique, to obtain a measure of " unit potency ". Again, no difference could be detected between the large and small strains. Selection for weight had thus brought about a correlated change in the weight of the pituitary gland, but without affecting its unit potency. Edwards found also that a similar conclusion applied to the rate of thyroid secretion.

It seems a reasonable hypothesis at this stage that the difference in ovulation rate between Falconer's large and small strains is probably due to correlated changes in organ sizes, rather than to any basic changes in physiological or endocrinal relationships. However, the evidence for this statement is by no means conclusive.

Rahnefeld, Boylan and Comstock (1962) estimated the genetic correlation between post-weaning growth and litter size. The value that they found, $+0 \cdot 153$, was not significantly different from zero. Their conclusion was that the correlated response, if any, of litter size to selection for growth should be in the direction of larger litters. Qualitatively, at least, this conclusion agrees with previous work on selection for weight.

Roberts (1961) examined the fertility of large and small mice from a different point of view. He measured their lifetime production of offspring, and found that the large strains had a drastically reduced length of reproductive life. On average, they produced $4\frac{1}{2}$ litters over their lifetime, against 11 or so in the small strains. On account of this, the small strains eventually weaned almost twice as many offspring as the large strains, in terms of lifetime production.

Roberts also examined the weights of the large and small strains over their lifetime. He found that the differences in mean weights, established by selection at six weeks of age, became magnified as the animals grew older, though the proportionate difference remained fairly stable. The two large strains showed one interesting common feature, and one important difference. They were similar in that, once their maximum weight had been achieved, both strains showed a marked decline in weight, probably caused by a depletion of body fat reserves. The two strains were different in that, though they both achieved the same maximum weight, one strain reached this maximum at six months of age, while the other did not reach it until one year old. This established a genetic difference in the shape of the growth curve, when plotted against age, and thus indicates some genetic independence between maximum weight and the prior rate of growth.

A relative difference between Goodale's and MacArthur's large mice in their degree of fatness was noted earlier. It is nevertheless a matter of observation that mice selected for large body weight usually become very fat. Fowler (1958), working with Falconer's strains, made a quantitative study of the carcass composition of mice selected for large and small body sizes. Compared with small mice, and compared also with the one control strain available for study, the large strains developed more fat, as a percentage of carcass weight, with correspondingly less protein and associated water. These differences had, in most cases, become established by six weeks, which was the age at selection, and the trend increased with increasing age. By twelve weeks of age, the percentages of carcass fat in the two large strains studied by Fowler were $27 \cdot 6\%$ and $15 \cdot 0\%$, while the corresponding figures for the small strains were $16 \cdot 3\%$ and $9 \cdot 7\%$. This difference between the large strains was statistically significant.

In a further study, Fowler (1962) found that the gross efficiency of food utilisation was higher in large than in small mice, between three and five weeks of age. Over this period, the gross efficiency declined, and thereafter the strains were indistinguishable. This decline was presumably due to an increase in maintenance costs in comparison with the weight gained. A parallel analysis in terms of energetic efficiency revealed the superiority of the large strains up to four weeks of age, when the growth rate was high, and after six weeks of age, when fat was being deposited at an increased rate. In between these ages, the strains were similar. Thus, Fowler established that selection for weight may bring about genetic differences in efficiency of food conversion. Fowler showed further that the greater efficiency of the large mice could not be attributed entirely to a greater proportion of ingested food being absorbed from the gut; the large mice absorbed also a greater proportion of protein. This evidence suggests a genetic basis of the efficiency of digestion.

The association of large body weight with a higher percentage of fat in the carcass was the subject of a study by Hull (1960). Fowler, in her 1958 paper, had suggested, reasonably enough, that " if selection had been made at 5 weeks of age when fat deposition has barely commenced, the differences in weight between animals might largely be attributable to differences in protein and associated water, and heritable differences in fatness might therefore have been excluded ". Hull's study constituted a test of this hypothesis. He selected for high body weight in three separate lines at the ages of 3, 4½ and 6 weeks. He followed the correlated changes in the percentage of fat in the carcass, and, contrary to all expectations, found this three-week line to be unmistakably the fattest, while the line selected at 4½ weeks was fatter than the one selected at six weeks. Hull speculated that at the age of three weeks, fat deposition had just begun in some animals, which were therefore selected and would go on to lay down large amounts of fat. However, this argument should apply *a forteriori* to his line selected at 4½ weeks. Hull's results have therefore not been fully explained.

The studies quoted here provide ample evidence that selection for body weight in the mouse may result in profound correlated changes in other characters. To the extent that mouse characters have analogues in domestic livestock, the direction of some of these correlated changes is not always favourable. For instance, a decline in reproductive longevity, or in carcass quality, might constitute some economic disadvantage. Some corrective procedures may therefore be required if intense and prolonged selection is practised for rapid growth.

The non-additive genetic variance of body weight

Most of the material in the preceding sections has been concerned, in the language of quantitative genetics, with various aspects of the additive genetic variance of body weight in the mouse. The results of some workers would suggest that non-additive genetic variance, with respect to this trait, does not even exist. Thus, Chai (1956*b*) concluded—" the dominance effect contributed to the total variability, if any, was considered to be trivial ". More recently, Miller, Legates and Cockerham (1963), in a detailed study involving the impressive total of 2879 mice, found no evidence of any non-additive hereditary variance with respect to three- and six-week weights. Other workers, *e.g.* Falconer (1953), have found directional dominance towards large size. Some of the recent literature on body size has been directed towards the exploration, and possible exploitation, of any non-additive variance that may exist. An attempt is given here to summarise the main conclusions.

An unexpected result was reported by Mason, Nicholson, Bogart and Krueger (1960). They crossed four inbred strains according to a diallel scheme and examined various characters in the parental strains and the crosses. Considering only their results on body weight for present purposes, the only significant departures of cross means from mid-parental values that they observed were all in the direction of *reduced* growth in the hybrids. For 45-day weight, this was true for two of six possible crosses, and for growth between 21 and 45 days, a significant negative deviation was observed for four of the six. The authors interpret these results in terms of a conflict at the physiological level—in that an inbred mother may not be able to meet all the requirements of her cross-bred offspring. The same authors (1957) had reported that one of the crosses included in their later report suffered a severe, inherited vitamin deficiency if reared by one of the parental inbreds. Negative heterosis was also observed in one of five crosses by Franks, Fechheimer and Cohen (1962), in all aspects of growth except birth weight. Oddly, perhaps, this instance was observed in a cross between two sublines of the C3H strain. However, in all their other crosses, Franks *et al.* found either positive heterosis or none at all. Furthermore, the number of cases that exhibited heterosis varied with the part of the growth curve that was studied. This paper perhaps reflects the general situation with respect to heterosis for growth in crosses between inbred strains of mice, namely, that it is a variable and unpredictable phenomenon.

Butler (1958) found that the amount of heterosis for weight, in a cross between two inbred strains, increased with age. At 60 days of age there was ample evidence that the F_1 was heavier, but the differences were not, generally speaking, significant at 30 days. Butler concluded that 30-day weight was largely a maternal character, and that an F_1 genotype thus conferred little advantage on growth rate when reared by an inbred mother. However, the same F_1 genotypes, when themselves used as mothers, increased the 30-day weight of their offspring by some 10 to 20%.

Butler also inbred MacArthur's large and small strains through 20 generations of brother-sister mating. Many of the lines that he started failed to survive the inbreeding, but, among the survivors, the large mice became smaller when inbred, while the small ones became larger. The decrease in

the large mice Butler attributes to inbreeding depression; the increase in the small mice is explained in terms of differential fertility, as the smallest of his small mice failed to reproduce under the additional burden of increasing homozygosity, and some lines therefore became extinct.

Carmon (1963) mated four strains, only one of which was highly inbred, in a complete diallel, and all his crosses showed considerable heterosis for weight both at 21 and 45 days of age. Carmon analysed his data also in terms of general combining ability, which is the average performance of a line in hybrid combination, and specific combining ability, which measures whether specific crosses deviate from expectation based on the average performance of the parental lines. While he found abundant evidence of general combining ability, Carmon's mean squares for specific combining ability were trivially small in his analysis of body weights.

If Carmon's finding is of general application to body weight in the mouse, then it has implications for certain selection schemes designed to exploit specific combining ability. Further evidence that this approach may not be profitable is supplied by Comstock, Singh and Enfield (1963). They crossed a strain selected for increased growth, described by Rahnefeld *et al.* (1963) and quoted earlier, to a long inbred line at each generation of selection, to measure the incidental effect of the selection on combining ability. The superior growth of the crosses, over the inbred level, did not differ significantly at any stage from half the increase observed in the selected line. Thus, general combining ability, associated with additive genetic variance, accounts for all the facts. The authors conclude that their results provided no evidence of non-additivity in genes affecting growth, though a previous observation of hybrid vigour on crossing the progenitors of the selected line (*see* Rahnefeld *et al.*, 1963) had indicated some non-additivity.

In view of these findings, it is not surprising that Hansson and Lindkvist (1962) could report no progress under a scheme of recurrent selection. Recurrent selection is designed to increase the specific combining ability by selecting animals purely on the basis of their crossing performance with an inbred tester. Hansson and Lindkvist's conclusion, with respect to body weight and other characters that they examined in their mice, was that recurrent selection, in combination with rotational crossbreeding, was at least not superior to what they termed " conventional " breeding, where selection was based on progeny testing.

In like manner, Newman (1960) made no progress with a reciprocal recurrent selection programme. Reciprocal recurrent selection differs from recurrent selection in that both parental stocks are selected on the performance of their crossbred progeny, when the parental stocks are mated together. Newman took two strains of mice that had been selected to, or near to, the limit for body weight. Thus the additive variance was largely exhausted, and reciprocal recurrent selection was applied in an attempt to capitalise on any non-additive variance that may be left. The method failed to yield any improvement over five cycles of selection.

This brief survey of the experimental evidence hence fails to yield a definitive picture of the non-additive genetic variance of body weight in the mouse. Heterosis in strain crosses is a common, but not an invariable, feature of the data. This is true even of those studies that allow for maternal effects, from variation in litter size or from other causes. The failure of the two experiments on recurrent and reciprocal recurrent selection, respectively must be read in the context of a general doubt about the efficacy of these methods. With the reservation that the situation has not yet been explored very extensively, we are forced to the conclusion that non-additive genetic variance does not figure prominently in the genetic architecture of body weight in the mouse.

Maternal effects on body weight

Though this review is written from a strictly genetical viewpoint, no discussion of body weight in the mouse is complete without a brief reference to maternal effects, as these have profound implications for the statistical analysis and biological interpretation of the character. Their sources may be many, as the growth of young mice is affected both by the uterine environment and by the nursing ability of the mother. An example of the latter effect was supplied by the cross-fostering experiment of Butler and Metrakos (1950). Thus, any variation in litter size, for example, leads to variation in weight from this cause alone. Whenever strains of mice are crossed reciprocally, the reciprocals reflect differences of the maternal environment in their growth. As examples of this phenomenon, several of the studies quoted in the previous section—Butler (1958), Carmon (1963), Franks *et al.* (1962), Mason *et al.* (1960), to name but a few—all draw attention to this feature of their data. Chai (1956*a*) notes that, in his data, maternal effects accounted for more than a quarter of the total variation in body weight at 60 days. Chai was examining crosses between strains which included Goodale's large and MacArthur's small mice, so he may have struck

an extreme example. But it is an impressive demonstration of the potential power of maternal effects.

A systematic study of the influence of maternal effects, as they affect growth in the mouse, was described by Brumby (1960). Using Falconer's (1953, 1955) large and small strains, and an unselected control strain, Brumby employed the techniques of ova transplantation and fostering to separate prenatal from postnatal maternal effects, having first satisfied himself that neither technique, of itself, affected the weight of the mouse at any stage. Brumby concluded that selection for body size had resulted in different maternal environments, which can therefore be rated as an additional correlated response. Further, this correlated response was asymmetrical; whereas the large strain barely equalled the control in maternal performance, the small strain showed a marked deterioration. A major portion of this maternal influence was exerted prenatally, though postnatal maternal influences were also evident. The postnatal performance, especially, was susceptible to modification according to the genetic constitution of the young that were suckled. For instance, while small strain females proved to be of equal lactational performance to the large strain females when rearing young of the small strain, they were markedly inferior when rearing young from the large strain.

Brumby also found one effect that could be interpreted as cytoplasmic inheritance. Reciprocal crosses between the large and small strains were gestated and suckled by females from the unselected control. These reciprocals differed in weight at all ages. After careful exclusion of sex linkage, Brumby suggested the possibility, curious though it may seem, that the cytoplasm of the small strain enhances body size to a greater degree than does the cytoplasm of the large strain. However, Brumby also notes that at the time of their transfer, eggs from the small strain were at a later stage of development than those from the large strain, which may have been sufficient to account for the difference.

There is therefore abundant evidence for maternal effects in body weight, and Brumby found that such effects persisted at least until the mouse was three months old. It was mentioned earlier that in many studies with the mouse, especially those of Falconer and his associates, complications due to maternal effects are avoided by using a within-family method of selection. However, Bateman (1963) rightly asks whether individual selection, or especially between-family selection, could not bring about heritable changes due to genes that act through a maternal effect that is independent of weight. To test this, Bateman mass selected mice for five-week weight, though two-thirds of the superiority of the selected animals could be ascribed to differences between families. Bateman's experimental procedures and his reasoning are too involved to explain in detail here, but he concludes that, of the maternal effects featured in his data, one-quarter must be ascribed to the direct result of maternal weight, while the remainder reflects other aspects of the maternal genotype.

The general conclusion from this brief examination of maternal effects is that they are a prominent feature of data on the weight of mice, and that any departure from a within-family method of selection may well magnify their effects. In mouse experiments, the tendency has been to try to avoid maternal effects not directly caused by weight, in order to simplify the genetic interpretation of the data. However, the objectives in livestock improvement may be different, and appropriate methods of selection may have to be chosen accordingly.

Implications of studies on the body weight of the mouse for animal breeding

This survey of the literature on the body weight of the mouse has revealed a genetic situation that is primarily additive in nature, and largely uncomplicated by interactions either at the genetic level or with the environment. Selection has usually been effective in bringing about marked changes in weight, and the limits to selection are usually not reached for twenty generations or more. To the extent that these results may be generalised to the larger animals, weight changes may thus be brought about easily, if required. Furthermore, the genetic architecture of the character is such that economical selection schemes, for instance those based on performance testing, may have a greater efficiency than the slower and more expensive progeny tests. Performance testing, of course, has already become standard practice with domestic livestock selected for weight. The possible dangers are, firstly, that selection in one environment may not be entirely appropriate to other environments, and secondly, that correlated changes, for instance, in productivity or in carcass quality, may not always be favourable. The indication from mouse studies would be that these correlated changes should be followed carefully as selection for increased growth proceeds.

REFERENCES

[*A.B.A.* = *Animal Breeding Abstracts*]

BAKELS, F. 1963. Untersuchungen an Töchtern männlicher Albinomäuse in zwei Fütterungsgruppen. (Ein Beitrag zum Problem der Erbe-Umwelt-Interaktion.) *Arb. Inst. Tierz. Vererb.- u. KonstForsch. Univ. München*, No. 5: 32 pp. [*A.B.A.*, **32**, No. 2264.]

BARNETT, S. A. 1965. Adaptation of mice to cold. *Biol. Rev.*, **40**: 5-51. [*A.B.A.*, **33**, No. 2715.]

BATEMAN, N. 1963. Selection and a maternal effect. In *Genetics today. Proc. XIth int. Congr. Genet.* [*The Hague*], 1963, Vol. I (Abstr.): 252. [*A.B.A.*, **32**, No. 422.]

BOHREN, B. B., HILL, W. G., and ROBERTSON, Alan. 1965. Some observations on asymmetrical correlated responses to selection. *Genet. Res.* [*Camb.*] (In the press.)

BRUMBY, P. J. 1960. The influence of the maternal environment on growth in mice. *Heredity*, **14**: 1-18. [*A.B.A.*, **28**, No. 2136.]

BUTLER, L. 1952. A study of size inheritance in the house mouse. II. Analysis of five preliminary crosses. *Canad. J. Zool.*, **30**: 154-171. [*A.B.A.*, **20**, No. 1178.]

BUTLER, L. 1958. The inheritance of litter size, body weight, and variability, in a cross between two inbred strains of mice. *Canad. J. Zool.*, **36**: 969-983. [*A.B.A.*, **27**, No. 950.]

BUTLER, L., and METRAKOS, J. D. 1950. A study of size inheritance in the house mouse. I. The effect of milk source. *Canad. J. Res., D*, **28**: 16-34. [*A.B.A.*, **18**, No. 647.]

CARMON, J. L. 1963. Heterosis, combining ability, and maternal effects in mice. *J. Genet.*, **58**: 225-231. [*A.B.A.*, **31**, No. 3074.]

CHAI, C. K. 1956*a*. Analysis of quantitative inheritance of body size in mice. I. Hybridization and maternal influence. *Genetics*, **41**: 157-164. [*A.B.A.*, **24**, No. 1796.]

CHAI, C. K. 1956*b*. Analysis of quantitative inheritance of body size in mice. II. Gene action and segregation. *Genetics*, **41**: 165-178. [*A.B.A.*, **24**, No. 1797.]

CHAPMAN, A. B. 1951. Effectiveness of selection in laboratory animals. *J. Anim. Sci.*, **10**: 3-8. (Symp. 1949 Meet. Amer. Soc. Anim. Prod.)

CHAPMAN, A. B. 1961. Contributions of laboratory animals to research in livestock improvement. In *Germ plasm resources. Publ. Amer. Ass. Advanc. Sci.*, No. 66: 225-238.

COCKREM, F. 1959. Selection for relationships opposite to those predicted by the genetic correlation between two traits in the house mouse (*Mus musculus*). *Nature* [*Lond.*], **183**: 342-343. [*A.B.A.*, **27**, No. 953.]

COCKREM, F. 1963. The effects of temperature on two strains of mice of different body weight and tail length. In *Genetics today. Proc. XIth int. Congr. Genet.* [*The Hague*], 1963, Vol. I (Abstr.): 154. [*A.B.A.*, **32**, No. 411.]

COMSTOCK, R. E., SINGH, M., and ENFIELD, F. D. 1963. Selection for growth in mice: crossbred performance. (Abstr.) *J. Anim. Sci.*, **22**: 1109. (Midwest. Sect. Amer. Soc. Anim. Sci.)

CROWCROFT, P., and ROWE, F. P. 1961. The weights of wild house mice (*Mus musculus* L.) living in confined colonies. *Proc. zool. Soc. Lond.*, **136**: 177-185. [*A.B.A.*, **29**, No. 1592.]

DALTON, D. C., and BYWATER, T. L. 1963. The effect of selection for litter size and litter weight at weaning in mice maintained on two diets. *Anim. Prod.*, **5**: 317-326. [*A.B.A.*, **32**, No. 1330.]

EDWARDS, R. G. 1962. The size and endocrine activity of the pituitary in mice selected for large or small body-size. *Genet. Res.* [*Camb.*], **3**: 428-443. [*A.B.A.*, **31**, No. 1413.]

FALCONER, D. S. 1952. The problem of environment and selection. *Amer. Nat.*, **86**: 293-298. [*A.B.A.*, **23**, No. 940.]

FALCONER, D. S. 1953. Selection for large and small size in mice. *J. Genet.*, **51**: 470-501. [*A.B.A.*, **21**, No. 1843.]

FALCONER, D. S. 1954. Validity of the theory of genetic correlation. An experimental test with mice. *J. Hered.*, **45**: 42-44. [*A.B.A.*, **22**, No. 1642.]

FALCONER, D. S. 1955. Patterns of response in selection experiments with mice. *Cold Spr. Harb. Symp. quant. Biol.*, **20**: 178-196. [*A.B.A.*, **25**, No. 1396.]

FALCONER, D. S. 1960*a*. Introduction to quantitative genetics. *Edinburgh and London: Oliver & Boyd Ltd.* ix+365 pp.

FALCONER, D. S. 1960*b*. Selection of mice for growth on high and low planes of nutrition. *Genet. Res.* [*Camb.*], **1**: 91-113. [*A.B.A.*, **28**, No. 1522.]

FALCONER, D. S., and KING, J. W. B. 1953. A study of selection limits in the mouse. *J. Genet.*, **51**: 561-581. [*A.B.A.*, **21**, No. 1844.]

FALCONER, D. S., and LATYSZEWSKI, M. 1952*a*. Selection for size in mice on high and low planes of nutrition. In *Quantitative inheritance. Ed. by E. C. R. Reeve and C. H. Waddington. London: H.M.S.O.* Pp. 145-151.

FALCONER, D. S., and LATYSZEWSKI, M. 1952*b*. The environment in relation to selection for size in mice. *J. Genet.*, **51**: 67-80. [*A.B.A.*, **21**, No. 332.]

FALCONER, D. S., and ROBERTSON, Alan. 1956. Selection for environmental variability of body size in mice. *Z. indukt. Abstamm.- u. VererbLehre*, **87**: 385-391. [*A.B.A.*, **24**, No. 1252.]

FOWLER, Ruth E. 1958. The growth and carcass composition of strains of mice selected for large and small body size. *J. agric. Sci.*, **51**: 137-148. [*A.B.A.*, **27**, No. 398.]

FOWLER, Ruth E. 1962. The efficiency of food utilization, digestibility of foodstuffs and energy expenditure of mice selected for large or small body size. *Genet. Res.* [*Camb.*], **3**: 51-68. [*A.B.A.*, **30**, No. 1960.]

FOWLER, Ruth E., and EDWARDS, R. G. 1960. The fertility of mice selected for large or small body size. *Genet. Res.* [*Camb.*], **1**: 393-407. [*A.B.A.*, **29**, No. 984.]

FRANKS, E., FECHHEIMER, N. S., and COHEN, C. 1962. An examination of heterosis in crosses of certain inbred strains of mice. *Ohio J. Sci.*, **62**: 178-184. [*A.B.A.*, **31**, No. 3075.]

GOODALE, H. D. 1938. A study of the inheritance of body weight in the albino mouse by selection. *J. Hered.*, **29**: 101-112. [*A.B.A.*, **6**: 125.]

GOODALE, H. D. 1941. Progress report on possibilities in progeny-test breeding. *Science*, **94**: 442-443. [*A.B.A.*, **10**: 45.]

HAMMOND, J. 1947. Animal breeding in relation to nutrition and environmental conditions. *Biol. Rev.*, **22**: 195-213. [*A.B.A.*, **15**: 227.]

HANSSON, A., and LINDKVIST, G. 1962. Selection for general and specific combining ability with regard to body weight in mice. *K. LantbrHögskol. Ann.* [*Uppsala*], **28**: 215-233. [*A.B.A.*, **31**, No. 2351.]

HARRISON, G. A. 1963. Temperature adaptation as evidenced by growth of mice. *Fed. Proc.*, **22**: 691-697. Discussion: 697-698. (Proc. int. Symp. on temperature acclimation.) [*A.B.A.*, **32**, No. 414.]

HULL, P. 1960. Genetic relations between carcass fat and body weight in mice. *J. agric. Sci.*, **55**: 317-321. [*A.B.A.*, **29**, No. 988.]

KING, J. W. B. 1950. Pygmy, a dwarfing gene in the house mouse. *J. Hered.*, **41**: 249-252. [*A.B.A.*, **19**, No. 790.]

KING, J. W. B. 1955. Observations on the mutant "pygmy" in the house mouse. *J. Genet.*, **53**: 487-497. [*A.B.A.*, **24**, No. 286.]

KORKMAN, N. 1957. Selection with regard to the sex difference of body weight in mice. *Hereditas* [*Lund*], **43**: 665-678. [*A.B.A.*, **26**, No. 935.]

KORKMAN, N. 1961. Selection for size in mice in different nutritional environments. *Hereditas* [*Lund*], **47**: 342-356. [*A.B.A.*, **30**, No. 1969.]

LEWIS, W. L., and WARWICK, E. J. 1953. Effectiveness of selection for body weight in mice from inbred and outbred populations derived from common parent stocks. *J. Hered.*, **44**: 233-238. [*A.B.A.*, **22**, No. 671.]

MACARTHUR, J. W. 1944*a*. Genetics of body size and related characters. I. Selecting small and large races of the laboratory mouse. *Amer. Nat.*, **78**: 142-157. [*A.B.A.*, **12**: 209.]

MACARTHUR, J. W. 1944*b*. Genetics of body size and related characters. II. Satellite characters associated with body size in mice. *Amer. Nat.*, **78**: 224-237. [*A.B.A.*, **12**: 209.]

MACARTHUR, J. W. 1949. Selection for small and large body size in the house mouse. *Genetics*, **34**: 194-209. [*A.B.A.*, **17**, No. 1007.]

MACARTHUR, J. W., and CHIASSON, L. P. 1945. Relative growth in races of mice produced by selection. *Growth*, **9**: 303-315. [*A.B.A.*, **16**, No. 616.]

MCNUTT, W., and DILL, R. E. 1963. Strain differences in mice on a high salt diet. Genetic and physiological comparisons between two strains. *J. Hered.*, **54**: 297-303. [*A.B.A.*, **32**, No. 2284.]

MASON, R. W., NICHOLSON, H., BOGART, R., and KRUEGER, H. 1957. Inherited differences in vitamin requirements among mice. *J. Pharmacol.*, **119**: 165-166.

MASON, R. W., NICHOLSON, H. H., BOGART, R., and KRUEGER, H. 1960. Predominance of hybrid losses or negative heterosis in mouse crosses. In *Biometrical genetics. Ed. by O. Kempthorne. Un. int. Sci. biol., B (Colloq.)*, No. 38: 188-209.

MILLER, R. H., LEGATES, J. E., and COCKERHAM, C. C. 1963. Estimation of nonadditive hereditary variance in traits of mice. *Genetics*, **48**: 177-188. [*A.B.A.*, **31**, No. 3079.]

NEWMAN, J. A. 1960. Reciprocal recurrent selection for body weight in the mouse. *Ph.D. Thesis, Univ. Edinburgh.*

RAHNEFELD, G. W., BOYLAN, W. J., and COMSTOCK, R. E. 1962. Genetic correlation between growth rate and litter size in mice. *Canad. J. Genet. Cytol.*, **4**: 289-295. [*A.B.A.*, **31**, No. 1428.]

RAHNEFELD, G. W., BOYLAN, W. J., COMSTOCK, R. E., and SINGH, M. 1963. Mass selection for post-weaning growth in mice. *Genetics*, **48**: 1567-1583. [*A.B.A.*, **32**, No. 1357.]

ROBERTS, R. C. 1961. The lifetime growth and reproduction of selected strains of mice. *Heredity*, **16**: 369-381. [*A.B.A.*, **30**, No. 1239.]

TAYLOR, St. C. S. 1954. A study of directional dominance in quantitative inheritance. *Ph.D. Thesis, Univ. Edinburgh.*

WARWICK, E. J., and LEWIS, W. L. 1954*a*. Growth and reproductive rates of mice produced by intercrossing inbreds and outbreds of three size groups. *J. Hered.*, **45**: 35-38. [*A.B.A.*, **22**, No. 1672.]

WARWICK, E. J., and LEWIS, W. L. 1954*b*. Increase in frequency of a deleterious recessive gene in mice from a population under selection. *J. Hered.*, **45**: 143-144. [*A.B.A.*, **22**, No. 1673.]

YOUNG, G. B. 1953. A study of genotype-environment interaction in mice. *J. agric. Sci.*, **43**: 218-222. [*A.B.A.*, **21**, No. 1404.]

PAPER 24

Some contributions of the laboratory mouse to animal breeding research

Part II: Fertility.

Animal Breeding Abstracts, 33 (4), 515-526. 1965

by

R.C. ROBERTS

SOME CONTRIBUTIONS OF THE LABORATORY MOUSE TO ANIMAL BREEDING RESEARCH

PART II

R. C. Roberts

A.R.C. Unit of Animal Genetics, Institute of Animal Genetics, Edinburgh 9

II. Genetic Analysis of Litter Size and Fertility in the Mouse

Data on some aspects of fertility, in particular the size of the litter, are easily gathered for any colony of laboratory mice. Indeed, such data are usually available as a by-product of the recording system. However, fertility is patently a highly complex character. If, as is often the case, fertility is defined operationally as " the size of the litter when it is first examined " (probably a few hours after birth), then this is a comprehensive measurement that has been determined sequentially by ovulation rate, fertilisation, zygotic survival, implantation, embryonic viability, parturition and neonatal survival of the young. Litter size can therefore be variously affected by the genotype of the dam and of the litter itself, or indeed even by the genotype of the sire. In addition, the environmental sources of variation are manifold. The implication of this complexity is that data on fertility must be interpreted with particular care; any differences found may have a multiplicity of causes, or at the other extreme may be attributable to but one of the processes indicated above.

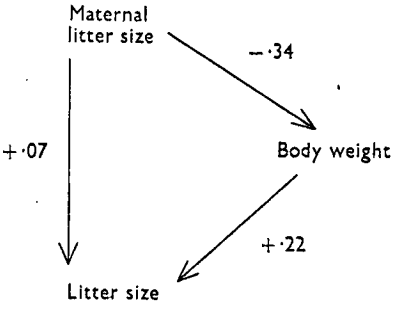
This review will be centred on experiments that have been designed specifically to elucidate the genetic control of fertility in the laboratory mouse. Litter size at birth is the commonest criterion of fertility, though, where adequate data exist, some of the components of litter size will be examined separately.

More so than for most characters, litter size has the practical advantage for experimentation of being easily measured. It is also fortunate that the distribution of litter size around its mean is usually sufficiently close to normal for the usual statistical tests to be valid without transformation of scale. But, apart from convenience, fertility is a popular character for experimentation because of its intrinsic importance. As in the case of farm animals, the fertility of a stock of mice affects the cost of its maintenance and the ease of its management. Furthermore, it bears directly on the rate of progress that may be expected from any programme of genetic improvement.

Maternal effects on litter size

It is convenient at this juncture to discuss briefly one important non-genetic source of variation in litter size, as some understanding of its effect is essential to the genetic interpretation of experimental data. As early as 1929, MacDowell, Allen and MacDowell noted that larger mice tend to produce larger litters, and conversely, the smaller the mouse, the smaller, on average, is her litter. There is a well-known maternal effect on the size of the litter associated with the mother's body size. From this, some complications arise. A large mouse tends to produce a large litter, with the result that the individual weights of her daughters are depressed. These daughters still reflect this handicap when they are mated, and thus they tend to produce, in turn, smaller litters, as a consequence of being themselves reared in a large litter. The net effect would thus be a negative regression of litter size on the size of the litter in which the mother was born, unless there also exists a positive genetic pathway, expected of a heritable character. These interactions were placed on a quantitative basis by Falconer (1955), who calculated standardised partial regression

coefficients relating litter size to the body weight of the dam and the size of the litter in which the dam was born. His results are summarised in the figure:



As expected, the body weight of a dam is negatively correlated with the size of the litter in which she was born, but is positively correlated with the size of the litter which she produces. The product of the two partial regression coefficients is −0·07, which would be the regression of litter size on maternal litter size if no other pathway were operative. However, the direct genetic pathway, measured as the partial regression of litter size on maternal litter size, holding body weight constant, is +0·07. Thus litter size is affected by two pathways of equal magnitude but opposite sign, which explains why the direct regression of litter size on maternal litter size, when measured, is zero (Falconer, 1955).

Falconer (1964) extended his treatment of maternal effects. This more sophisticated analysis yielded a direct estimate of the maternal effect relating the litter size of a mouse to the size of the litter in which that mouse was born. This estimate agreed well with that obtained from the earlier study, from which Falconer was able to conclude that the influence through body weight was an adequate explanation of the observed maternal effect on litter size.

Though the maternal effect on litter size is relatively small in absolute terms, it is sufficient, as Falconer explains, to account for some otherwise serious discrepancies between various estimates of the genetic parameters of litter size, and between the predictions to which those parameters lead.

Selection for litter size

In the preceding section, it was seen that if body weight is held constant, the regression of daughter's litter size on mother's litter size becomes positive. Falconer's two methods of estimating this pathway of genetic transmission yielded regression coefficients of +0·07 and +0·10, corresponding to heritabilities of 14 and 20% respectively. Thus, if the maternal effect can be circumvented, selection for litter size ought to give some response. An experiment designed to achieve this effect has been described by Falconer (1955, 1960c, 1963, 1964), the separate publications referring to different stages and to various aspects of his study. There were two selected lines, one for increased and the other for reduced litter size, and an unselected control line. The complications due to maternal effects were avoided by selecting among groups of full-sister (litter-mate) females, each group being mated to the one male. One female was selected from each group either for high or for low litter size, according to the direction of the selection. Sterile matings were, of course, discounted. Thus, the method of selection was for female reproductive performance within a family of full-sisters, and no selection was applied to any contribution of the males to the litters that they sired. Each family contributed equally to succeeding generations. A group of full-sister females was mated to a male from another family to minimise any irregularities that might otherwise result from inbreeding. Inbreeding, nevertheless, did accumulate at the rate of roughly 1% per generation, as estimated from the control line, on account of the small number of parents per generation. Nevertheless, despite this accumulated inbreeding, litter size remained stable in the control line over 31 generations. We shall return to this fact in the context of inbreeding.

The number of mice reared per litter was not standardised by artificial augmentation or reduction. This procedure, of course, reintroduced the maternal effect; mice in the high line were reared in large litters, which placed them at an environmental disadvantage, while mice in the low

line were favoured by being reared in small litters. Thus the maternal effect opposed the direction of selection, but this was a constant effect and did not complicate the selection procedure. The effect was, nevertheless, conspicuous in the first generation, as the response in both lines was markedly opposite to the direction of the selection. But thereafter the responses were in the expected direction, except for the irregular fluctuations expected of any selection programme based on fairly small numbers.

Both selected lines responded more or less linearly over 20 generations or so and then remained stable for a further 11 generations. At the limit, the high line averaged 9·2 mice per litter and the low line 6·0, compared to 7·6 in the control. The response was therefore symmetrical in absolute terms. It is, however, more meaningful to evaluate the response in terms of the selection differential, which was greater in the high line. As a result, the realised heritability was found to be 8% for upward selection and 23% for downward selection, while the figure calculated from the divergence between high and low assumed an intermediate value of 13%. When these values, based on within-family selection, were converted to the basis of individual heritabilities, they became 15, 40 and 22%, respectively, for the high line, low line and divergence. In these terms, the response to selection was markedly asymmetrical in the high and low lines.

The final divergence of 3·2 mice per litter between the two selected lines represents a difference of 1·6 times the original phenotypic standard deviation, and 3·3 times the additive genetic standard deviation. Thus the total response on this basis is very small compared to the responses obtained for body weight reported in the first part of this review.

The asymmetry of the response to selection is possibly explained, at least in part, by the qualitative differences in the nature of the response, described by Falconer in his 1963 paper. When the components of litter size were examined separately, it was found that the progress in the high line could be adequately attributed to an increase in the ovulation rate. But, curiously, the ovulation rate in the low line had also increased compared to the control. The reduced litter size in the low line was attributable entirely to a marked increase in early post-implantational death among the foetuses. Crosses between the selected lines and the control showed that the embryonic mortality in the low line was attributable to the mothers and not to the embryos themselves. Falconer suggested that this property of the low line females might be due to recessive genes, initially at a low frequency, which would increase in frequency rapidly under selection for low litter size. The effect of any reduction in their frequency by selecting for large litters would scarcely be noticeable, even if effective. But the increase in ovulation rate in the low line is more difficult to explain, even though we assume that genes causing prenatal losses largely masked any genetic variation in ovulation rate.

Bateman (in press) reports another selection experiment for litter size using a stock derived from Falconer's and differing only in the manner of selection. In his high line, Bateman selected parents from the largest litters, and in the low line from the smallest litters. This procedure—between-family selection—differs radically from Falconer's within-family method. Bateman did not keep a control line. After 12 generations of selection, the high line had a mean litter size at birth of 11·1, against 5·5 in the low line. When the components of litter size were examined, Bateman found that his high and low lines differed in ovulation rate, implantation rate, foetal survival and possibly neonatal viability. The low line was particularly prone to failure of implantation. The differences between the two lines were all characteristics of the strain of the dam, and Bateman showed that neither male fertility nor the genetic constitution of the offspring had contributed to the response to selection. Bateman therefore broadly confirms Falconer's main results, though differences appear in some of the details.

The only other selection experiment for litter size in the mouse appears to be that reported by Dalton and Bywater (1963). They selected for high litter size at weaning, both on a normal diet and on a diluted diet. In neither case did the selected line deviate significantly from the unselected control strains over 14 generations of selection. It is possible, though surprising if true, that litter size at weaning time has a much lower heritability than litter size at birth. This would explain the difference between the results obtained by Falconer and by Bateman, on the one hand, and by Dalton and Bywater on the other.

The general conclusion from selection studies for litter size must be that the trait responds to selection at birth, while it may not respond to selection at weaning time. As yet, we have insufficient evidence to estimate the expected magnitude of the response for selection at birth. Whereas Falconer reported a modest response, Bateman obtained a much larger divergence in fewer generations. As both Falconer's and Bateman's stocks had a common origin, we may

speculate that the different systems of mating employed may have contributed to the difference between the two experiments. The full-sib correlation for litter size is reported by Falconer (1964) to be only 0·107. It would appear on theoretical grounds (*see* Falconer, 1960*a*) that this correlation is too low to warrant a within-family method of selection.

Litter size is an obvious, and probably a major, component of natural fitness in a polytocous animal such as the mouse. The dogma of quantitative genetics is that such traits should not exhibit much additive genetic variance. Yet, Falconer in his study was able to derive individual heritabilities of up to 40% in his stocks, and this is perhaps slightly surprising. One possible explanation is that litter size, especially the size of the first litter, is not such a major component of fitness as we think. Roberts (1961) showed that mean litter size need not necessarily be correlated with the total number of young that a mouse may produce over its lifetime. But probably a better explanation of Falconer's relatively high heritabilities is that litter size may have an intermediate optimum, in that the very large litters may give rise to fewer adult mice than smaller litters. Falconer (1960*c*) states that the number of mice weaned is reduced once the number born alive exceeds 13 or so. If some such situation has existed during the evolutionary history of the laboratory mouse, then we have a mechanism for the retention of a fair amount of additive genetic variance in litter size, or at least in some of its components.

Crosses between inbred lines

Many genetical studies on litter size have taken the form of crosses between inbred strains. As a general point, we may note that this approach is somewhat circumscribed by the peculiar genetic constitution of inbreds. Apart from their high level of homozygosity, inbred strains by their very nature cannot carry any lethal genes and are unlikely to carry any seriously deleterious genes either. Likewise, of course, crosses between inbred strains do not carry such genes. Further, a population of strain crosses differs from an outbred population in that no genes are at a low frequency. For instance, if we take a four-strain cross, then no allele segregating in the derived population can have a frequency lower than 0·25. When we add to all this the fact that an inbred strain can produce only one kind of gamete, and a cross between four strains therefore represents only four gametes from a hypothetical outbred population, accidents of sampling can play a decisive part when we synthesise a segregating population from inbred strains. We can thus appreciate that for a normally outbreeding organism such as the mouse, work with inbred strains has to be undertaken at a considerable loss of generality. This is not necessarily a criticism of all work involving inbred strains; for some studies they are eminently suitable. The point is that any strain cross represents a unique and peculiar situation, and we should not be disturbed by any apparent anomaly between work based on inbred strains and studies carried out on other inbreds or, in particular, work with outbred populations. Studies on inbred strains should nevertheless be examined for the information that they do contain; any difference between strains kept under uniform conditions indicates genetic variation.

The increase in the fertility of crossbred mice has attracted attention for a long time. Castle (1926) was aware of the phenomenon. Fortuyn (1932) crossed two strains of albino mice differing in fertility and found that both strains reared a larger number of offspring when the litter was hybrid, compared to the level of the pure strains. Grüneberg (1939) noted that the fertility of some crossbred mice exceeded any that he had encountered elsewhere. Thus the evidence seemed to be accumulating that crossbreeding increased fertility.

A more comprehensive study was reported by Eaton (1941, 1953). He gathered nine inbred strains and made crosses to test their fertility. About half the crosses produced F_1 litters exceeding those of either parent strain, though some crosses were even inferior. The combination of three inbred strains, using a hybrid dam, gave a greater increase in litter size. Eaton proceeded to form new inbreds from crosses between some of his initial strains. His new strains seem to have survived the inbreeding for a further six or seven generations. But when these new inbreds were in turn crossed, the effect of heterozygosity in the litter was much reduced compared to the previous study. The conclusion from these studies is that the genotype of both the dam and of the litter itself may affect litter size, but that the former is of greater relative importance. However, Forsthoefel (1954) found that the genotype of the litter had a conspicuous effect on fertility. He took one inbred strain (BALB/c), and split some litters. Some females were mated to their brothers, while their litter-mates were mated to inbred males of another strain. The effect of the crossbreeding was to increase litter size from 4·8 to 6·8, the statistical significance of the difference

being beyond doubt. Forsthoefel suggested that the increase was brought about by the masking of recessives that reduced viability in the uterus.

It is probably safe, for all practical purposes, to neglect any influence of the male on litter size, other than through the effect on the genotype of the litter, although the evidence on this point is slightly equivocal. Both Falconer (1955) and Bateman (1965) found that males had no direct influence on litter size when mated at random to groups of dams. On the other hand, Finn (1964), in a similar situation, found a statistically significant effect. Finn notes, however, that embryonic mortality probably contributed to the effect. But it is possible that males may occasionally have a low fertilising capacity, as Krzanowska (1960*b*) found in one inbred line. However, as far as most of the variation is concerned, litter size is partly determined by the genetic constitution of the dam and partly by the genetic constitution of the litter itself. More evidence on this dual determination of litter size was provided by Butler (1958). He took two inbred strains (C57 and BALB) and crossed them reciprocally. One strain (BALB) showed an increase in litter size when bearing a crossbred litter, whereas the other did not. Butler was able to exclude the possibility that this effect was due to the low fertilising capacity of BALB males, for these males did not reduce the litter size of other more prolific females. The fertility of the F_1 females significantly exceeded that of either parental strain. The general summary of Butler's study is that crossbreeding in the dam had a greater and more uniform effect in increasing fertility than crossbreeding in the litter.

However, crossing does not always result in increased fertility. A contrasting situation is depicted by Bogart, Mason, Nicholson and Krueger (1958), and by the same authors (Mason *et al.*, 1960) in another publication. Their material consisted of four strains that had been maintained as closed colonies for many generations and which they crossed according to a diallel scheme. Heterosis in respect of litter size in the first cross was obtained in only three of the twelve possible crosses (treating reciprocals as being different). In five crosses the litter size was near the mid-parental value. In the remaining four crosses litter size was reduced, markedly so in two of them. The authors interpret their results in terms of endocrine function and note also the highly specific vitamin requirements of some crossbreds which the pure-strain mothers were unable to provide.

The variability of the possible effects of strain crossing on fertility is illustrated further by Franks, Fechheimer and Cohen (1962). They crossed an inbred strain with five others. In four of the crosses, heterosis in respect of litter size was noted. The remaining cross, however, had a litter size inferior to that of either parent.

Other studies involving strain crosses have concentrated on the effect of crossbreeding on various components of litter size. Lyon (1959), in an investigation of the mutational load in three inbred strains, crossed them in all possible combinations. She found that post-implantational mortality in the crossbred progeny was, on average, 9·6% less than in the pure strains. Her strains differed widely in the reduction of mortality that they showed on crossing, the individual percentages being 4·8, 7·6 and 16·9 for the CBA, C3H, and 101 strains, respectively. Lyon suggests that some, though not all, of the mortality within inbred strains was due to recessive lethals arising by mutation in these strains.

McCarthy (1965) reports a study designed specifically to measure the effect of crossing inbred strains of mice on litter size and to determine the potential litter size at various stages of gestation. He crossed four strains in all possible combinations. Three of these strains (CBA, C57 and R111) showed an increase in litter size on crossing, whereas the fourth (JU) did not. It is interesting to note that McCarthy, like Lyon, found that the CBA strain showed only a slight increase on crossing. The increases that McCarthy found were all attributable to the effect of crossing on one stage of gestation, namely early post-implantational mortality. McCarthy does not favour the hypothesis that new recessive lethal mutations cause embryonic mortality within strains and prefers to interpret his data in terms of heterosis in embryonic viability.

A similar experiment is described by Martin, Harrington and Hill (1963). Unfortunately, this report appears only in abstract form, where the designations of the strains are not given, so that direct comparison with Lyon and with McCarthy is not possible. Martin *et al.* likewise crossed four strains in a complete diallel. Unlike the other two studies quoted, they found that the implantation rate was higher when the litter was crossbred, although differences in the initial ovulation rate and in the final litter size at birth were not statistically significant. When the crossbred mice were used as mothers, the ovulation rate was still the same as in the purebred mice, but thereafter the crossbreds showed a higher litter size at all stages of gestation. Martin *et al.* noted that in their

pure line matings, only 36% of the embryonic mortality occurred after implantation. When the litters were crossbred, 58% of the mortality occurred at this stage, and when the crossbreds were employed as parents, the figure rose to 100%.

It was noted at the beginning of this section that differences between experiments utilising inbred strains should cause no surprise. Such differences have now been amply demonstrated. Nevertheless, the general conclusion from studies on inbred strains is fairly clear. Crossbreeding in mice usually, though not always, causes an increase in litter size. Further, this increase is usually, though again not always, associated to a greater extent with crossbreeding in the dam rather than with crossbreeding in the litter. Heterosis in litter size, being present more often than not, clearly suggests some non-additive genetic variation in litter size in the mouse.

Non-additive genetic variation in litter size

The studies quoted in the preceding section in almost every case afford a qualitative reflection of some non-additive genetic variance in litter size. In this section we shall attempt to evaluate the importance of this variance quantitatively. There are two slightly different approaches to this question, although the two are closely related and entirely interdependent. Either the total variance is partitioned empirically into general and specific combining abilities, reflecting the additive and non-additive components respectively, or else the variance is partitioned formally into its components according to standard statistical procedures. If the data were sufficiently detailed, either approach could be adopted at will.

Martin *et al.* (1963), quoted previously, analysed their F_1 data and reported that, in their material, specific combining ability was a more important source of variation than general combining ability. This agrees with the finding of Miller, Legates and Cockerham (1963), whose analysis yielded an estimate of non-additive genetic variance representing 28% of the total phenotypic variance, or 58% of the total genetic variance, thus leaving only 42% of the genetic variance attributable to additive sources. In contrast, Carmon (1963), who mated three outbred stocks and one inbred in a complete diallel, found that his estimates of general and specific combining abilities were both trivially small and did not even approach the magnitude required for statistical significance.

Carmon's results are concordant with an earlier attempt by Bowman (1962) to exploit any non-additive genetic variance in litter size by means of recurrent selection. His experiment was a sire selection programme, where each sire was assessed on the mean litter size of a half-sib group of daughters produced by mating the sires each to a number of dams from an inbred tester. Though in the experiment mean litter size did increase somewhat, the increase was no greater than the expected response if all the variance between sires in crossing performance was additive genetic variance. Yet in the initial generation the hybrid progeny were not intermediate between the selected stock and the inbred tester, indicating some dominance. Bowman was forced to conclude that either recurrent selection is an inefficient way of exploiting non-additive variance, or else that there was no non-additive variance (apart from some dominance) in his material.

The situation that has emerged is reminiscent of that pertaining to body weight, discussed in the first part of this review. Whereas we have considerable evidence of non-additivity in litter size from crossbreeding work, this often seems to be attributable to a general heterotic factor that does not always yield measurable amounts of non-additive genetic variance. To this we must add the reservation that there is a paucity of detailed analyses on which to base any general conclusion.

Effect of inbreeding on the litter size of outbred populations

The deleterious effects of inbreeding on the reproductive performance of an outbreeding organism are well-known, and the laboratory mouse is no exception. The highly inbred strains of mice commonly available in mouse laboratories today represent the relatively few chance survivors from the hundreds (possibly thousands) of attempts to establish inbred strains in the past. Therefore, there has been much opportunity for selection for fertility and viability during the formation of the strains, and the question arises whether the increases in fertility shown in strain crosses is to be attributed to this selection. Roberts (1960) argues that random differentiation of gene frequencies during inbreeding is alone insufficient to account for the higher fertility of crosses, without selection at some stage. He reports an experiment calculated to test the influence of natural selection during inbreeding on the fertility of subsequent crosses. From an outbred

population he derived 30 partially inbred strains and crossed these in a random manner when the inbreeding coefficient had reached 0·50. No selection was applied during the inbreeding stage, except for natural selection operating within lines. Litter size declined markedly during the inbreeding at the rate of about half a mouse per 10% inbreeding. When the lines were crossed, the effect of crossbreeding in the litter was to increase litter size somewhat. Only when the crossbreds were themselves employed as parents was fertility restored to the outbred level. The conclusion was that natural selection operating during the inbreeding did not cause any increase in litter size.

A complementary study to this was reported by Bowman and Falconer (1960). As natural selection within lines does not operate to increase fertility, Bowman and Falconer investigated the influence of (*a*) artificial selection for large litters during inbreeding, and (*b*) selection between lines, lines becoming extinct as they became too infertile to maintain. From the same base population as in Roberts' (1960) study, twenty inbred lines were derived and maintained by full-sib mating. A decline in litter size of 0·56 mouse per 10% inbreeding was observed, a figure in close agreement with that obtained by Roberts. In ten of the lines there was deliberate selection for large litters, but their litter sizes declined at exactly the same rate as in the other ten, in which there was no selection. This answers the first question asked. By the time the inbreeding coefficient reached 76%, only three of the original 20 lines were still surviving, and, of these, only one survived indefinitely. The individual histories of the three survivors showed that, initially, each had a litter size below the mean of the population, but that this litter size did not decline as inbreeding progressed.

When the inbreeding coefficient in the three surviving lines had reached 81%, they were crossed. The litters produced by the crossbred progeny were larger than those of the outbred control by about two mice per litter. Thus, selection between lines during inbreeding resulted in a substantial increase in litter size. This raised the interesting possibility that repeated inbreeding and crossing might result in further progress. However, Bowman and Falconer reported negative results for two such further cycles. In terms of animal improvement, therefore, this method offers but limited scope.

An inbreeding depression in litter size of roughly 0·5 mouse per litter per generation was noted both by Roberts and by Bowman and Falconer, as a result of rapid inbreeding by full-sib mating. This situation contrasts strongly with that found by Falconer (1960c) in an outbred control stock. Here, the inbreeding coefficient accumulated slowly through a restriction on the population size. Falconer notes that over 31 generations the inbreeding coefficient rose to 0·32, but that there was no evidence of any reduction in litter size over this period. It seems therefore that slow inbreeding, though cumulative, need not have the effects that we normally associate with the mating of close relatives. It is possible that natural selection may counteract the effects of a gradual rise in the inbreeding coefficient to maintain litter size at the normal level.

The effect of heterozygosity, in the dam and in the litter, on litter size shows an anomaly in Roberts' (1960) study, quoted earlier. In the first generation of inbreeding, when the dams were still outbred, litter size fell by 1·39 as a result of raising the inbreeding coefficient of the young to 0·25. In the next generation, when the dams were inbred as well, the decline was no greater than that expected as a result of further inbreeding in the young. However, on crossing, when the inbreeding coefficient was changed from 0·5 to 0, the effect of the increased heterozygosity in the young was to raise litter size by a mere 0·51, while the effect on the dams was a further improvement of 2·27 over the previous generation. The relative importance of inbreeding on the dam and in the litter therefore is not consistent. Roberts interprets his results as a maternal limitation on litter size in inbred mothers, irrespective of the heterozygosity of the young.

In his first cross (inbred dams, crossbred litters), Roberts found no evidence of either general or specific combining ability (*cf.* Carmon (1963), quoted earlier). In the second cross, when the dams were crossbred as well, there was some general combining ability but still no suggestion of any specific combining ability. However, the point is made that an inbreeding coefficient of only 0·50 largely precludes the possibility of observing any specific combining ability in strain crosses. Roberts concludes that, as higher coefficients are not readily attainable in mammals, inbreeding followed by crossing is not an efficient method of animal improvement.

Falconer and Roberts (1960) examined the stage of gestation at which litter size was reduced by inbreeding. Ovulation rates, implantation sites and the number of live embryos were estimated by dissection of pregnant females, both inbred (50-60%) and outbred, from the same stock. There was no reduction in ovulation rate as a result of the inbreeding, but the reduction in the number of implants in inbred females was sufficient to account for their smaller litters. However, Falconer

and Roberts' finding, though consistent throughout their data, may not be a general one. Krzanowska (1960*a*) reports that for one inbred strain, litter size was reduced at all stages by inbreeding, while another inbred line suffered considerable loss of ova through lack of fertilisation. However, Krzanowska's two strains had a history of inbreeding extending over 20 generations, which renders the comparison with an outbred control somewhat tenuous.

While there is not the slightest doubt that litter size in outbred populations of mice shows a considerable decline on inbreeding, the possibility remains that in different stocks different stages of gestation may be affected. If this is so, then the situation is consistent with that pertaining to the effects of crossing on the components of litter size discussed in an earlier section.

Implications of studies on litter size in the mouse for animal breeding

This examination of the genetics of litter size in the mouse yields a slightly paradoxical conclusion. Fertility is probably the favourite example of a trait closely connected to natural fitness, where little additive genetic variance is expected, and where non-additive genetic variance is deemed to figure prominently. But whereas inbreeding studies have conclusively established the presence of directional dominance, other evidence of non-additive variance is less convincing. Indeed, methods of exploiting any non-additive variance to improve litter size have so far proved inferior to more straightforward methods of selection. One would therefore hesitate to advocate costly, laborious and time-consuming schemes, such as cyclical inbreeding and crossing or recurrent selection, as methods of animal improvement. We may perhaps venture to suggest that the success story of hybrid corn may not easily be repeated in mammals. Or, stated another way, perhaps ordinary selection may yet prove to be a better way of improving corn.

III. Studies on the Genetic Control of Sex Ratio in the Mouse

The control of sex ratio in farm animals has often been the subject of experimentation and much discussion because of its obvious genetic and economic implications. It seems safe to assume that the sex ratio of roughly 50:50 observed in higher organisms is the product of natural selection for an optimal value, which suggests the possibility of some genetic variation in the ratio. Although the mechanism of sex determination at fertilisation in mammals leads to a primary sex ratio of 50:50, this nevertheless can result in differences between species in the secondary sex ratio at birth (Nalbandov, 1964). Further, within one species, the mouse, Howard, McLaren, Michie and Sander (1955) reported significant heterogeneity between the sex ratios of the six different F_1s from four inbred strains. Such variation is clearly attributable to genetic sources.

However, Falconer (1954) reported a selection experiment for sex ratio in mice, and a completely negative result was obtained. Falconer had two selected lines, one selected for a high proportion of males and the other for a high proportion of females. Over four generations the proportion of males increased slightly in both lines, but in neither case did the increase even approach statistical significance. Falconer reports further that in a wide search among his mouse stocks, he found no evidence whatever of any heterogeneity of sex ratio between families; this result indicates the absence of any genetic variation in sex ratio. Falconer adds the caution that as the measurement of sex ratio by its nature involves a large error variance, any real variation in sex ratio may thereby be obscured.

In contrast to Falconer's deliberate attempt to modify the sex ratio, Weir (1953) found that he had brought about marked differences in the sex ratio of two lines of mice as a concomitant deviation to selection for blood *p*H. The effect was immediately apparent in the first generation and became firmly established by a subsequent nine generations of inbreeding. His high-*p*H line showed a change from 50% males in the unselected population to 55%, whereas the low-*p*H line had 47% males. In a further publication, Weir (1955) showed that the characteristic sex ratio of each line was determined wholly by the male parent.

However, attempts to modify the sex ratio by selecting for blood *p*H have not yielded uniform results. Wolfe (1961) reports a partial repeat of Weir's experiment, conducted in the same laboratory. Whereas Weir had used venous blood, Wolfe selected for *p*H levels in the arterial blood. While the blood-*p*H level again responded rapidly to selection, the correlated response in sex ratio was in the opposite direction to that observed by Weir. In Wolfe's high-*p*H line, the percentage of males dropped to 46, while it rose to 52·5 in his low line.

The general conclusion is therefore that, although there is no doubt about the individual results

of Weir and of Wolfe, no reliable method has yet been found that modifies the genetic determination of sex ratio in a predictable fashion.

IV. Studies on the Genetic Control of Lactation in the Mouse

The usefulness of the laboratory mouse for experimental studies on lactation is severely restricted by the impossibility of obtaining a direct measurement of milk yield. Attempts to circumvent the problem of measurement have been based on the employment of the pre-weaning weights of litters of mice, to reflect the lactational performance of the mother. In a preliminary study, Falconer (1947) suggested that the 12-day total weight of the litter, expressed as a percentage of the mean weight of litters of that size, might be a satisfactory measurement of lactation.

The sources of variation in Falconer's measurement, and its validity as a measure of lactation, are discussed by Bateman (1954, 1957). Bateman shows that the measurement is indeed largely a characteristic of the mother and is but little affected by the properties of the young. However, the maternal influence is partly prenatal in origin. Whereas 73% of the variation in total litter weight at 12 days is maternally determined, only 32% is associated with postnatal growth. Bateman therefore suggests that the measurement should be termed " maternal performance ", as it is obviously only in part a reflection of lactation. Bateman shows further that the maximum amount of milk is available when the litter is born, and that the supply is rapidly and permanently cut down to the requirements of the litter. An association between body weight and maternal performance is established, but age has no independent effect on the measurement, once its influence through body weight has been taken into account.

Bateman selected on the total 12-day weight of the litter, as a measure of maternal performance in mice. Though the full results of this experiment have, unfortunately, not been published, some account of Bateman's results is given by Falconer (1955) in a review of selection responses. Briefly, the selection was both for high and for low litter weights, where the number in the litter was standardised. The experiment was conducted mostly on a mass selection procedure, though some of the early generations were selected in a much more complicated way, which need not concern us here. The responses had some extraordinary features. Upward selection over twenty-one generations yielded but slow and irregular progress. However, at all stages of the experiment, reversed selection from the high line for low litter weights had an immediate and pronounced response. Downward selection showed that this rapid response continued for about six generations and then ceased. However, when upward selection was renewed from the low line, at its limit, a large response was repeatedly observed for high litter weights. The realised heritability from the divergence was always about 0·50, though the asymmetry of the responses at different mean levels implies a much higher value for downward selection from the high line, and for upward from the low. Obviously, as is pointed out, the interpretation of these results cannot be based on simple premises. Natural selection appears to be a potent factor in preventing fixation at both high and low levels. It is also highly probable that only a few genes control the major part of the response.

Some Japanese work has been aimed towards a more direct measurement of mammary function in the mouse. The technique, described by Yoshida (1961), is to inject immature mice at weaning with synthetic oestrogen, and to measure the area of the excised mammary glands at 35 days. The measurement obtained is referred to as the " mammary growth response ", and was reported by Yoshida to have a heritability of 0·59. Nagai, Yoshida and Naito (1962) report that selection for the mammary growth response had been effective for 15 generations. Body weight showed some increase over this period, suggesting some genetic association between the two traits. However, more information about the connection between " mammary growth response " and lactation is required if the measurement is to become useful in the genetic analysis of milk yield, using the mouse as the experimental animal. In the meantime, the measurement of " maternal performance " employed by Falconer and Bateman, though perhaps not closely connected with lactation, may be of greater relevance to studies with farm animals. But, obviously, until the lactation of a mouse can be more adequately measured, its genetic control cannot be fully explored.

General Conclusion

This review has been concentrated on the genetic investigation of certain traits in the mouse that may bear some analogy to situations encountered by the breeders of domestic livestock. To a certain extent, these studies were probably motivated by heuristic objectives, and the context of

their theoretical stimulation may deviate considerably from the more pragmatic issues of animal improvement. Any information that they contain of relevance to animal breeding may therefore be quite fortuitous. The corollary of this is that some information of potential interest to the animal breeder may not have been pursued very extensively. We must also reiterate the caution that extrapolation from one species to another may lack complete validity. Indeed, even a rapid perusal of the mouse literature is sufficient to establish that the repeatability of results within the species is something less than absolute.

From this, it would be easy to argue that any generalisation from the mouse laboratory to the farm should be withheld. Nevertheless, to my mind, one conclusion emerges clearly. A major portion of animal breeding effort is directed towards the improvement of growth and fertility, in that order. It is almost an article of faith that evolution has characterised these two traits by different genetic moulds. Yet the mouse information is entirely self-consistent in one respect, and it applies to both growth and fertility. It shows that there is no case on record, in an extensive literature, where ordinary selection, or a minor variant thereof, has been surpassed by elaborate and costly schemes designed to exploit non-additive genetic variance. The profitable utilisation of such variance still remains a challenge for the future. Other species, of course, may yield a different experience. But for the present, it would seem a safe maxim that breeding schemes should not neglect the potential of some standard form of selection, unless a departure is indicated by supplementary information for the species concerned. At the very least, there should be some cognisance of the fact that only in special circumstances will the non-additive genetic variance be likely to exceed the additive.

Many questions of concern to the animal breeder cannot yet be answered from experimental work with the mouse. There is as yet insufficient information on the repeatability of responses to selection, and on the limits to these responses. The efficacy of various selection schemes has not been explored in enough detail. There is little knowledge, for any mammal, of the effect of population size on selection responses, nor is there any precise knowledge on the balance between the intensity of selection and the counteracting effects of inbreeding.

It is probable that some of the mouse experiments of the future will be directed towards these questions, while they will continue, of course, to provide didactic and dialectical exercises for some research students who aspire to contribute to the breeding of better livestock.

REFERENCES

[*A.B.A.* = *Animal Breeding Abstracts*]

BATEMAN, N. 1954. The measurement of milk production of mice through preweaning growth of suckling young. *Physiol. Zoöl.*, **27**: 163-173. [*A.B.A.*, **23**, No. 803.]

BATEMAN, N. 1957. Some physiological aspects of lactation in mice. *J. agric. Sci.*, **49**: 60-77. [*A.B.A.*, **25**, No. 2009.]

BATEMAN, N. 1966. Ovulation and post-ovulation losses in strains of mice selected for large and small litter sizes. *Genet. Res.* [*Camb.*]. (In the press.)

BOGART, R., MASON, R. W., NICHOLSON, H., and KRUEGER, H. 1958. Genetic aspects of fertility in mice. *Int. J. Fertil.*, **3**: 86-104. [*A.B.A.*, **28**, No. 834.]

BOWMAN, J. C. 1962. Recurrent selection. II. An experimental study with mice and *Drosophila*. *Genet. Res.* [*Camb.*], **3**:333-351. [*A.B.A.*, **31**, No. 1602.]

BOWMAN, J. C., and FALCONER, D. S. 1960. Inbreeding depression and heterosis of litter size in mice. *Genet. Res.* [*Camb.*], **1**: 262-274. [*A.B.A.*, **29**, No. 384.]

BUTLER, L. 1958. The inheritance of litter size, body weight, and variability, in a cross between two inbred strains of mice. *Canad. J. Zool.*, **36**: 969-983. [*A.B.A.*, **27**, No. 950.]

CARMON, J. L. 1963. Heterosis, combining ability, and maternal effects in mice. *J. Genet.*, **58**: 225-231. [*A.B.A.*, **31**, No. 3074.]

CASTLE, W. E. 1926. The explanation of hybrid vigor. *Proc. nat. Acad. Sci.* [*Wash.*], **12**: 16-19.

DALTON, D. C., and BYWATER, T. L. 1963. The effect of selection for litter size and litter weight at weaning in mice maintained on two diets. *Anim. Prod.*, **5**: 317-326. [*A.B.A.*, **32**, No. 1330.]

EATON, O. N. 1941. Crosses between inbred strains of mice. *J. Hered.*, **32**: 393-395. [*A.B.A.*, **10**: 44.]

EATON, O. N. 1953. Heterosis in the performance of mice. *Genetics*, **38**: 609-629. [*A.B.A.*, **22**, No. 1066.]

FALCONER, D. S. 1947. Milk production in mice. *J. agric. Sci.*, **37**: 224-235. [*A.B.A.*, **15**: 266.]

FALCONER, D. S. 1954. Selection for sex ratio in mice and *Drosophila*. *Amer. Nat.*, **88**: 385-397. [*A.B.A.*, **23**, No. 941.]

FALCONER, D. S. 1955. Patterns of response in selection experiments with mice. *Cold Spr. Harb. Symp. quant. Biol.*, **20**: 178-196. [*A.B.A.*, **25**, No. 1396.]

FALCONER, D. S. 1960*a*. Introduction to quantitative genetics. *Edinburgh and London: Oliver & Boyd Ltd.* ix+365 pp.

FALCONER, D. S. 1960*c*. The genetics of litter size in mice. *J. cell. comp. Physiol.*, **56** (Suppl. 1, *Symp. on mammalian genetics and reproduction*): 153-167. [*A.B.A.*, **29**, No. 1607.]

FALCONER, D. S. 1963. Qualitatively different responses to selection in opposite directions. In *Statistical genetics and plant breeding. Ed. by W. D. Hanson and H. F. Robinson. Publ. nat. Res. Coun. nat. Acad. Sci.* [*Wash.*], No. 982: 487-490.

FALCONER, D. S. 1964. Maternal effects and selection response. In *Genetics today. Proc. XIth int. Congr. Genet.* [*The Hague*], 1963, Vol. III: 763-774.

FALCONER, D. S., and ROBERTS, R. C. 1960. Effect of inbreeding on ovulation rate and foetal mortality in mice. *Genet. Res.* [*Camb.*], **1**: 422-430. [*A.B.A.*, **29**, No. 982.]

FINN, C. A. 1964. Influence of the male on litter size in mice. *J. Reprod. Fertil.*, **7**: 107-111. [*A.B.A.*, **32**, No. 2310.]

FORSTHOEFEL, P. F. 1954. A study of the effect of heterozygosis on litter size in the mouse, *Mus musculus*. *Ohio J. Sci.*, **54**: 135-141. [*A.B.A.*, **23**, No. 816.]

FORTUYN, A. B. D. 1932. A case of hybrid vigor in the albino *Mus musculus*. *Proc. Soc. exp. Biol.* [*N.Y.*], **29**: 784-786.

FRANKS, E., FECHHEIMER, N. S., and COHEN, C. 1962. An examination of heterosis in crosses of certain inbred strains of mice. *Ohio J. Sci.*, **62**: 178-184. [*A.B.A.*, **31**, No. 3075.]

GRÜNEBERG, H. 1939. Fertility in cross-bred mice. *J. Hered.*, **30**: 83-84. [*A.B.A.*, **7**: 237.]

HOWARD, A., MCLAREN, A., MICHIE, D., and SANDER, G. 1955. Genetic and environmental influences on the secondary sex ratio in mice. *J. Genet.*, **53**: 200-214. [*A.B.A.*, **23**, No. 1349.]

KRZANOWSKA, H. 1960*a*. Early embryonal mortality in inbred lines of mice and their crosses. *Bull. Soc. roy. belge Gynéc. Obstét.*, **30**: 719-728. [*A.B.A.*, **30**, No. 1970.]

KRZANOWSKA, H. 1960*b*. Studies on heterosis. II. Fertilisation rate in inbred lines of mice and their crosses. *Folia biol.* [*Kraków*], **8**: 269-279. [*A.B.A.*, **29**, No. 1613.]

LYON, M. F. 1959. Some evidence concerning the " mutational load " in inbred strains of mice. *Heredity*, **13**: 341-352. [*A.B.A.*, **28**, No. 309.]

MCCARTHY, J. C. 1965. The effect on litter size of crossing inbred strains of mice. *Genetics*, **51**: 217-222. [*A.B.A.*, **33**, No. 2734.]

MACDOWELL, E. C., ALLEN, E., and MACDOWELL, C. G. 1929. The relation of parity, age, and body weight to the number of corpora lutea in mice. *Anat. Rec.*, **41**: 267-272.

MARTIN, T. G., HARRINGTON, F. E., and HILL, D. L. 1963. Effects of mating system on ovulation, implantation and litter size in mice. In *Genetics today. Proc. XIth int. Congr. Genet.* [*The Hague*], 1963, Vol. I (Abstr.): 155. [*A.B.A.*, **32**, No. 451.]

MASON, R. W., NICHOLSON, H. H., BOGART, R., and KRUEGER, H. 1960. Predominance of hybrid losses or negative heterosis in mouse crosses. In *Biometrical genetics. Ed. by O. Kempthorne. Un. int. Sci. biol., B* (*Colloq.*), No. 38: 188-209.

MILLER, R. H., LEGATES, J. E., and COCKERHAM, C. C. 1963. Estimation of nonadditive hereditary variance in traits of mice. *Genetics*, **48**: 177-188. [*A.B.A.*, **31**, No. 3079.]

NAGAI, J., YOSHIDA, M., and NAITO, M. 1962. Establishment of mice strains with special reference to mammary growth response. VIII. Selection with inbreeding for the mammary growth response. *Jap. J. Breed.*, **12**: 246-254. [*A.B.A.*, **32**, No. 455.]

NALBANDOV, A. V. 1964. Reproductive physiology. Comparative reproductive physiology of domestic animals, laboratory animals, and man. *San Francisco and London: W. H. Freeman & Co.* 2nd ed. xiii+316 pp.

ROBERTS, R. C. 1960. The effects on litter size of crossing lines of mice inbred without selection. *Genet. Res.* [*Camb.*], **1**: 239-252. [*A.B.A.*, **29**, No. 406.]

ROBERTS, R. C. 1961. The lifetime growth and reproduction of selected strains of mice. *Heredity*, **16**: 369-381. [*A.B.A.*, **30**, No. 1239.]

WEIR, J. A. 1953. Association of blood-*p*H with sex ratio in mice. *J. Hered.*, **44**: 133-138. [*A.B.A.*, **22**, No. 271.]

WEIR, J. A. 1955. Male influence on sex ratio of offspring in high and low blood-*p*H lines of mice. *J. Hered.*, **46**: 277-283. [*A.B.A.*, **24**, No. 762.]

WOLFE, H. G. 1961. Selection for blood-*p*H in the house mouse. *Genetics*, **46**: 55-75. [*A.B.A.*, **29**, No. 1637.]

YOSHIDA, M. 1961. Effectiveness of selection with inbreeding on mammary growth response. I. Establishment of mouse strains different in mammary growth response by selection with inbreeding. *Jap. J. zootech. Sci.*, **32**: 125-133. [*A.B.A.*, **30**, No. 1244.]

PRINTED IN GREAT BRITAIN BY
OLIVER AND BOYD LTD.
EDINBURGH

PAPER 25

Some Concepts and Methods in Quantitative Genetics

Chapter 11, pp. 214-257. in Behaviour-Genetic Analysis. ed. Jerry Hirsch, McGraw Hill. 1967

by

R.C. ROBERTS

CHAPTER ELEVEN
SOME CONCEPTS AND METHODS IN QUANTITATIVE GENETICS[1]

R. C. Roberts

INTRODUCTION

Quantitative genetics is the study of the inheritance of characters where the individual effects of the genes concerned are obscured from view. Let us consider for a moment a typical example of such a character—body weight in the mouse. We know of several genes that affect body weight; for instance, the *yellow* coat-color gene makes the mouse noticeably heavier. Some mutant genes, such as *pygmy* or *obese*, are most easily identifiable through their effect on weight. However, the effect of such genes is large enough for them to be recognized on a given genetic background, and they are best studied through the established methods of formal genetics; their dominance, pleiotropic, and epistatic relationships can be described, and they can often be located on the linkage map of the mouse. But if we take a closer look at a litter of mice, segregating for, say, the *obese* mutant, we could immediately score the animals by eye into obese ones and "normals." Among the two classes, we could detect further differences in body weight, though perhaps we should require a balance to rank them accurately. We should discover further, if we were to carry out the appropriate experiments, that these smaller differences in weight are also in part heritable. But we can no longer recognize the genes

[1] *Acknowledgment*: Readers familiar with D. S. Falconer's *Introduction to Quantitative Genetics* will recognize the source of much of the material in this chapter. It is a pleasure to acknowledge my indebtedness to this source and to thank Dr. Falconer for reading this manuscript.

responsible for this variation, nor are we sure exactly in what way they affect body weight. Whereas we know that the *obese* gene increases the amount of fat, we could not be sure whether the residual variation in weight is due to slight variation in adiposity, in skeletal size, or in what. Through elaborate experimentation, we could perhaps discover answers to some of these questions, but we should fail to disentangle the individual genes involved because their effect is masked by other genes, perhaps having an identical effect. We are now in the domain of quantitative genetics, where individuals differ because they contain different samples of the alleles that potentially affect the character but where the investigator cannot know which particular alleles affect particular individuals. This kind of variation immediately implies a statistical approach. In practice, these *quantitative characters,* as we shall call them, are also subject to a certain amount of environmental variation. Our example of body weight is an obvious case; differences in nutrition, for instance, will obviously be reflected in the measurements. But environmental variation is not limited to quantitative characters, for it is to be found also in the expression of single genes.

It seems reasonable to presume, at this stage, that many aspects of behavior, such as activity and learning, require a quantitative scale of measurement in their treatment. Certainly, some genes exist whose behavioral effects are clear without refined measurements. As a specific example, the gene causing phenylketonuria in human beings has a well-known and marked effect on intelligence. This gene is therefore a source of variation in intelligence among individuals, but no one would argue that this gene, or genes like it, accounts for all the genetic variation in intelligence. Indeed, such genes do not appear as a cause of variation among the so-called "normal" individuals at all. Yet, these "normal" individuals are known to vary, and such variation is known to be, in part, of genetic origin. Similarly, phenylketonuric individuals vary among themselves, and it may be presumed that this variation also is caused, in part, by some of the genes that are responsible for variation among nonphenylketonuric individuals. Although genes such as the one causing phenylketonuria are dramatic in their effect, their occurrence is relatively infrequent, and they cannot therefore be regarded as an important source of variation in intelligence in a human population. The study of genetics is thus limited in scope if it is confined to genes causing "abnormalities." Quantitative genetics finds its application in the resolution of the "normal" range of variation.

The following, then, are the premises of quantitative genetics. A quantitative character is presumed to be affected by a number of genes, whose individual effects are small. Whereas it is clearly absurd to suppose that all the effects of such genes are interchangeable at the level of the primary gene action, the effects are nevertheless deemed to be interchangeable at the level of the measurement. This says no more than that a given body weight, for instance, can be the end product of several distinct developmental pathways. It is a further premise that the genes affecting a quantitative character can exhibit all the properties associated with the so-called "Mendelian" genes, such

as dominance, epistasis, and linkage, but that "quantitative" genes possess no properties that are unknown in formal genetics.

Quantitative genetics differs from formal genetics essentially by virtue of the fact that the effects of all the genes involved cannot be identified individually. It does not differ, as is sometimes supposed, because a character is measured rather than scored nor because such a character often exhibits a large amount of environmental variation. Superimposed on the framework of formal genetics is a statistical approach which has given rise to a specialized methodology. This has inevitably led to specialized concepts and to specialized thinking, but the concern is still with genes and not with mathematical relationships.

THE RECOGNITION OF THE GENETIC BASIS OF A QUANTITATIVE CHARACTER

Suppose we were interested in a character exhibiting a range of continuous variation—characters such as intelligence in a human population or activity scores in mice. We should need to know whether differences between individuals were in any way related to differences in their genotypes and also whether such genotypic differences occurred at one locus or at more. We should eventually need to know what proportion of the variation was attributable to genetic sources.

The first of these questions is sometimes easy to answer, especially if it is rephrased. Given known genotypic differences between animals, do corresponding differences appear in the character that we are examining? The answer to this question is facilitated in some laboratory species, such as mice or Drosophila, by the use of genetically homogeneous material, most commonly inbred strains. Any variation within such a strain is presumed to be environmental in origin. If, however, an analysis of variance reveals excess variation over this amount between strains kept under uniform conditions, there is clear evidence that the character shows genetic variation. It is preferable that several strains should be employed in the search for genetic variation in this way and that these should furthermore be of diverse origin. Even so, the absence of a difference between strains is not clear proof that the character shows no genetic variation in the species at large, for the strains examined may have been fixed for the same genes controlling this character or for genes whose sum effect is similar. However, this is perhaps unlikely. In species where inbred material is unavailable, it may be possible to substitute geographical races or even subspecies for inbred strains.

Having found genetic variation, we need to know whether it is controlled by one gene or by more than one. In the presence of environmental variation, this may not be as easy as it sounds. The approach is to examine for segregation groups of progeny from parents of known values, employing backcrosses, intercrosses, etc., exactly as we would if the character displayed little or no environmental variation, and where the classes of progeny might then be discrete. Perhaps the

most critical type of mating would be between parents of intermediate value, to see whether the full range of expression is obtained in the progeny, as expected on a single-locus model. It is seen that environmental variation would make it difficult to recognize more than one locus and well-nigh impossible to recognize even a single polyallelic locus. A confusing situation can also arise in the case of threshold characters. Such characters may under certain conditions mimic single-gene characters rather closely.

Unless we can establish that the genetic variation of the character is controlled by a single locus, or by a few recognizable loci, we have in effect a quantitative character, and the genetic questions to ask must be phrased accordingly. These questions will not refer particularly to individual animals, whose genetic constitution is unknown, but rather to populations of animals, where the sum effect of the genes is more clearly seen. We need to know how much of the total variation is due to various genetic causes, for it is axiomatic that the importance of a source of variation is proportional to the contribution it makes to the total variation. We need to know the extent to which relatives, on average, resemble one another, for this is what we really mean when we say that a character has a genetic basis. (All characters are genetic in the sense that genes are *sine qua non* for their existence.) We need to know whether we can manipulate the genes to change the level of expression of the character, otherwise our research work may become circumscribed. And we need to know how those genes are most readily manipulated, otherwise our time may be wasted. It is to questions such as these that much of what follows will be devoted.

THE SOURCES OF VARIATION IN THE POPULATION

Let us first examine what genetic factors affect the average performance of a population of animals, and what factors contribute to differences between individuals. To do this, we need to determine the *population mean* and the *genotypic variance*. An understanding of these parameters is necessary for the further development of ideas about quantitative genetics. The symbolism, terminology, and derivations used throughout this chapter follow those used by Falconer (1960).

The *phenotypic value* (P) of an individual consists of two parts, a *genotypic value* (G), determined by the individual's genetic constitution, and an *environmental deviation* (E), which may be either positive or negative. The environmental deviations are taken to be such that their sum over the whole population is zero.

$$P = G + E \qquad \text{where } \Sigma E = 0$$

It is a fundamental feature of the formulation that G and E are uncorrelated, and since some behaviorists may consider this an unwarranted premise, the point should perhaps be elaborated. It seems obvious that an animal's genotype may influence its choice of environ-

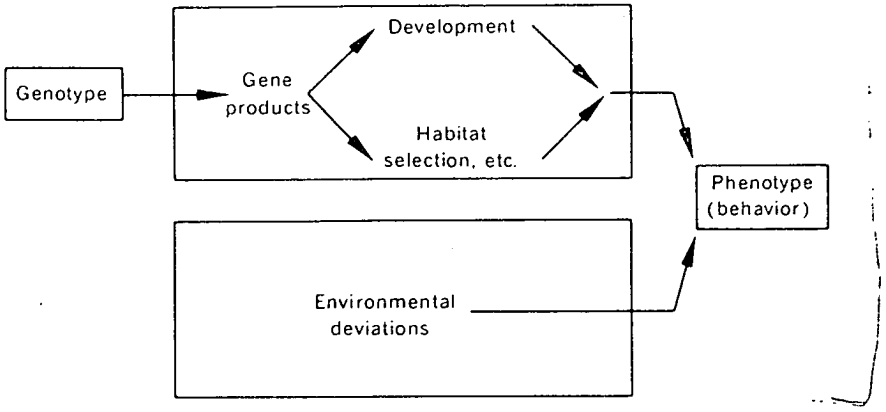


Figure 11.1 Relationship between genotype, environment, and phenotype, in schematic form, whereby the environmental deviations are *defined* as being independent of the genotype.

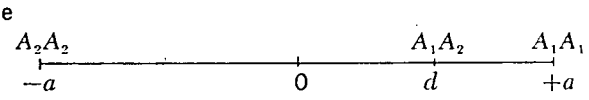
ment, which would therefore introduce a correlation between the two. This indeed may be so, even under a carefully regulated laboratory regime. However, this phenomenon, where it may exist, should be examined in the context of the pathway from the genotype to the phenotype, in this case behavior. The ramifications in this pathway are manifold and complex; for the present purposes, let us consider the grossly simplified version depicted in Figure 11.1. The genotype may influence the phenotype either by means of biochemical or other processes, labeled for convenience as "development," or by means of influencing the animal's choice of environment. But this second pathway, just as much as the first, is a genetic one; formally it matters not one whit whether the effects of the genes are mediated through the external environment or directly through, say, the ribosomes. In either case, the genotype affects the phenotype, and in this sense, all that comes between the two can be lumped together in a "black box" and treated as various parts of the same genetic process. This has complete operational validity, for the properties of a system can be explored (and often must be explored) without specifying completely the individual components of the system. There are of course powerful precedents for this approach in the physical sciences. But to return to the absence of a correlation between the genotypic values and the environmental deviations, the environment is *defined* as that which affects the phenotype independently of the genotype. If an effect stems from the genes, it is genetic; any other effect is an environmental one.

Since $\Sigma E = 0$, it follows that the average genotypic value of the individuals in a population is equal to their average phenotypic value and that either can be referred to as the population mean (M). Symbolically,

$$\overline{P} = \overline{G} = M$$

To see what genetic factors affect the mean level of performance in a population, we shall begin with a single locus with two alternative

alleles, A_1 and A_2. Let the frequency of A_1 be p, and that of A_2 be q, where $p + q = 1$. It follows that, if diploid organisms mate at random, there will appear among the offspring three genotypes, A_1A_1, A_1A_2, and A_2A_2 in the ratio of p^2: 2 pq : q^2, respectively. We must now assign scale values to the three genotypes. Let the homozygote A_1A_1 exceed in value the homozygote A_2A_2 by a quantity $2a$ on the scale of measurement. In other words, if the midpoint between the two homozygotes is regarded as the zero level, A_1A_1 has a deviation of $+a$, and A_2A_2 a deviation $-a$. Unlike many models of "Mendelian" dominance, the value of the heterozygote need not coincide with that of either homozygote. Let the deviation of the heterozygote, A_1A_2, from the zero level be d, which can assume either sign and any value.



Thus, if $d = +a$, this represents the complete dominance of the A_1 allele, with respect to this character, while if $d > +a$, it represents an *overdominant* situation. The illustration above depicts partial dominance of the A_1 allele.

We can now derive an expression for the population mean. The model can be expressed in tabular form as in Table 11.1. Multiplying each value by its frequency and summating over the three genotypes, the mean (m) of the population for this particular locus is

$$\begin{aligned} m &= p^2a + 2pqd - q^2a \\ &= a(p^2 - q^2) + 2dpq \\ &= a(p + q)(p - q) + 2dpq \\ &= a(p - q) + 2dpq \end{aligned}$$

since $p + q = 1$.

Summating over all relevant loci, the population mean (M) for the character measured is

$$M = \Sigma a(p - q) + 2\Sigma dpq$$

Let us consider briefly the factors that affect the mean genotypic value and thus the mean phenotypic value of the population. The mean depends on (1) the values of a and d and (2) the gene frequency q. For instance, it is easily seen that, at intermediate gene frequencies, the population mean is largely determined by the value of the heterozygotes, e.g., if $p = q = \frac{1}{2}$, $p - q = 0$ and the numerical value of pq is at a maximum. At more extreme gene frequencies, the heterozygotes

Table 11.1

Genotype	A_1A_1	A_1A_2	A_2A_2
Value	$+a$	d	$-a$
Frequency	p^2	$2pq$	q^2

have correspondingly less effect on the mean. It may be mentioned here that, since a and d are fixed values for any given locus under a specified set of conditions, the only genetic way of changing the population mean is by changing the gene frequencies.

The average genotypic value of a population is easily measured, as indicated, by the average phenotypic values. However, the genotypic values of individuals are theoretical values which could be measured only if a genotype could be replicated many times and placed in all the environments to which the population is subjected. In any case, the genotypic value of an individual, considered as a summation over many loci, is not of paramount interest in genetic work, for it refers to a unique combination of genes drawn at random from the population gene pool. Considering again at a single locus of a diploid organism, the two genes at that locus, whether the same or different, are represented in the gametes with equal frequencies. The part of the genotypic value of the parent transmitted to the offspring is the *average effect* of those two genes, considered separately. Take, for instance, the A_1 allele, uniting at random with other alleles of the A locus in the population. In p cases, it unites with another A_1 allele, giving genotypes of value $+a$; in q cases, it unites with A_2, giving genotypes of value d. The mean value of genotypes deriving from the A_1 alleles is therefore $pa + qd$. The average effect (α_1) of the A_1 allele is the deviation of these genotypes from the population mean for that locus:

$$\begin{aligned}\alpha_1 &= (pa + qd) - [a(p - q) + 2dpq] \\ &= q[a + d(q - p)]\end{aligned}$$

Similarly, it may be shown that the average effect of the A_2 allele is

$$\alpha_2 = -p[a + d(q - p)]$$

If we were now to allow one allele to be replaced by its alternative form, what would be the effect on the population mean? Suppose that an A_2 allele were to be replaced by an A_1. This hypothetical procedure enables us to determine the *average effect of gene substitution* at the locus in the population. The A_2 allele is distributed between the genotypes A_1A_2 and A_2A_2 in the ratio p/q. There is therefore a probability of p that the substitution would change the A_1A_2 genotype to A_1A_1 and thus change the value from d to $+a$. Likewise, there is a probability of q that the effect of the substitution would be to change the value from $-a$ to d. Thus, the average effect of gene substitution (α) is

$$\begin{aligned}\alpha &= p(a - d) + q(d + a) \\ &= a + d(q - p)\end{aligned}$$

The same formula can be derived by supposing that the substitution is in the opposite direction, i.e., if an A_1 allele were to be replaced by an A_2, except that the sign of the expression would then be negative. This expression shows that the average effect of a gene is, in fact, the average effect of a gene substitution, weighted by the opportunity

for that substitution to occur, i.e., by the frequency of the alternative allele. Thus,

$$\alpha_1 = q\alpha$$
$$\alpha_2 = -p\alpha$$

It is also seen that the average effect of a gene substitution is the difference between the average effect of the two alleles.

$$\alpha_1 - \alpha_2 = (q + p)\alpha = \alpha$$

These algebraic manipulations may appear for the moment to be somewhat irrelevant, but they are the foundations of some of the basic concepts of quantitative genetics. They lead immediately to one such concept, the *breeding value*, which, unlike the genotypic value, can be measured in an individual animal. It refers to the average effect of the individual genes that a parent transmits to its offspring, and we have just seen that this represents only a part of the genotypic value. For instance, in a fully dominant situation, where $d = +a$, the genotypic values of A_1A_1 and A_1A_2 are identical, whereas their breeding values are clearly different, because the latter will produce some A_2A_2 offspring while the former cannot. The breeding values of the three genotypes can be written as the sum of the average effect of their two genes considered separately. As the average effects were derived as deviations from the population mean, the breeding values are also expressed as deviations.

If an individual is mated to a sufficiently large number of animals drawn at random from the population, its breeding value can be measured directly as twice the deviation of the progeny group from the population mean. It has to be twice, because the gametes of the individual concerned unite at random with others, which by definition have a deviation of zero from the population mean. Thus, if a male mouse has progeny whose weight, on average, exceeds the population mean by 2.5 grams, the male's breeding value is +5 grams. If it is mated to a female whose own breeding value is also +5, the expected genotypic value of the offspring is +5. But if the female has a breeding value of −7, the expected genotypic value of the offspring is then −1 gram.

The breeding value is often called the *additive value* of the genotype, and the average effect of a gene may be referred to as its *additive effect*. The variance associated with this source will later be termed the additive variance. Although this terminology is probably self-explanatory, care should be taken not to confuse it with additive gene

Table 11.2

Genotype	Breeding value
A_1A_1	$2q\alpha$
A_1A_2	$(q-p)\alpha$
A_2A_2	$-2p\alpha$

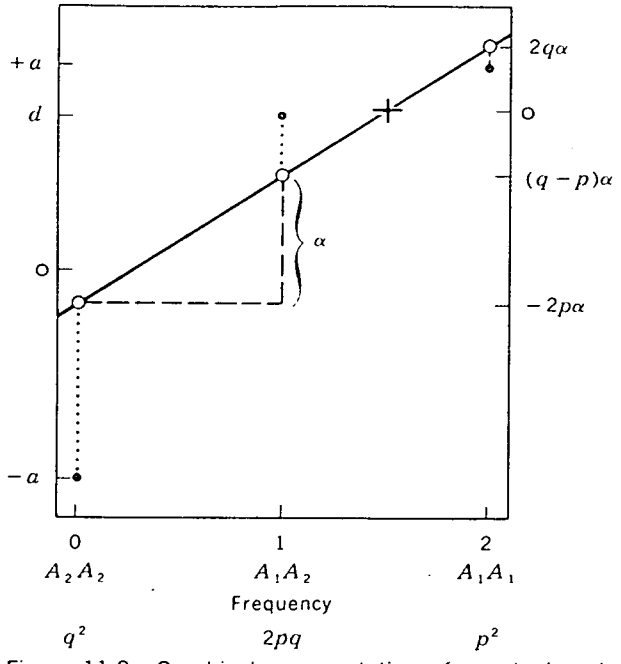


Figure 11.2 Graphical representation of genotypic values (closed circles) and breeding values (open circles) of the genotypes for a locus with two alleles A_1 and A_2, at frequencies p and q. Horizontal scale: number of A_1 genes in the genotype. Vertical scales of value: on left, arbitrary values as explained in the text; on right, deviations from the population mean. The figure is drawn to scale for the values $d = \frac{3}{4}a$ and $q = \frac{1}{4}$. (*From Falconer, 1960. By courtesy of author and publisher.*)

action. Additive gene action implies that the heterozygote value lies midway between that of the two homozygotes. The additive effect of a gene refers specifically to its average effect in the population, which has been shown to depend, *inter alia*, on the dominance. Thus, a fully dominant gene ($d = +a$) has a perfectly well-defined additive effect, while its action would be described as nonadditive.

The breeding or additive value, then, represents a part of the genotypic value. The difference between the two is represented graphically in Figure 11.2. This is a plot of values against the number of A_1 alleles. The closed circles represent genotypic values, while the line is the least-squares regression fit to these points, weighted by the number of individuals that each point represents. This weighting is therefore determined by the gene frequencies in that population. The value at which the line intersects the position of the genotypes on the gene-dosage axis represents the additive value of those genotypes (open circles). The difference between the additive value of the

heterozygote and that of either of the two homozygotes represents the additive effect of the gene substitution.

$$2q\alpha - (q - p)\alpha = (q + p)\alpha = \alpha$$

The deviation of each point from the line represents the part of the genotypic values not accounted for by the additive values. These deviations are known as *dominance deviations* and are always present when the degree of dominance is not zero. If we label the additive effect A and the dominance part D, then

$$G = A + D$$

and

$$P = A + D + E$$

In accordance with the terminology we have adopted, the dominance deviations may be referred to as *nonadditive* effects of the genes. The dominance deviations are the residual effects of combining genes in pairs, over and above the average effects of the genes considered separately.

The algebraic expressions for D for each locus are obtained by subtracting A from G, after first expressing G on the same basis as A, that is, as a deviation. Thus for the A_1A_1 genotype

$$\begin{aligned} G - M &= a - M \\ &= a - [a(p - q) + 2dpq] \\ &= 2qa - 2pqd \end{aligned}$$

Now,

$$\begin{aligned} A &= 2q\alpha \\ &= 2qa + 2q(q - p)d \end{aligned}$$

so that

$$\begin{aligned} D &= (G - M) - A \\ &= -2pqd - 2q(q - p)d \\ &= -2q^2d \end{aligned}$$

Repeating the procedure for the other two genotypes, we can derive Table 11.3.

This shows that only two factors contribute to the dominance deviations, the gene frequency and the heterozygote deviation. Whereas

Table 11.3

Genotype	A_1A_1	A_1A_2	A_2A_2
Frequency	p^2	$2pq$	q^2
Value	a	d	$-a$
$G - M$	$2q(a-pd)$	$a(q-p) + d(1-2pq)$	$-2p(a + qd)$
A	$2q\alpha$	$(q-p)\alpha$	$-2p\alpha$
D	$-2q^2d$	$2pqd$	$-2p^2d$

dominance, in the gene-action sense, contributes to α and thus to A, additively acting genes ($d = 0$) do not contribute to D. Figure 11.1 shows why the dominance deviations of the homozygotes are always negative if d is positive. It is worth emphasizing also that G, A, and D are all affected by the gene frequencies. This means that the values of these quantities for any locus refer specifically to one population and, strictly speaking, at one point in time. Any change in gene frequency will alter the relative proportion of heterozygotes to homozygotes. This will affect the slope of the line in Figure 11.2, which in turn alters everything.

For a single locus, the relationship $P = G + E = A + D + E$ is exhaustive; it fully describes the system, as far as it goes. But when we consider a genotype as a combination of several or many loci, another term must be added to accommodate the interaction between loci:

$$G = A + D + I$$

where I is the interaction deviation between two or more loci, representing the effect on the genotypic value of an individual over and above the average effects of those loci considered separately. This interaction can be of many different kinds, as in formal genetics. We need not consider the interaction term in any detail here, except to recognize it as a source of variation in the population.

Thus far, we have been concerned to identify the sources of variation in the population and to examine qualitatively the factors associated with each source that gives rise to variation. We must now study the actual variances a little more closely.

Since $P = G + E$, it follows that

$$V_P = V_G + V_E + 2\,cov_{GE}$$

where the Vs represent the variances associated with each source, and cov is the covariance. According to our definition of G and E, there is no covariance between the genotype and the environment. Thus,

$$V_P = V_G + V_E$$

In behavior, there may be psychometric problems to resolve, but once this is done, the phenotypic variance (V_P) is easily measured. However, it was said earlier that G (and therefore E) could not be measured for individual animals unless we could replicate the genotype at will and measure it under a range of environmental conditions. Now a situation that may approximate this, possibly very closely, is the case of a highly inbred line or of a cross between two such lines. In this case there is a replicated genotype, and the variance within such a genotype must be wholly environmental. In practice, it would obviously be preferable to obtain as many such genotypes as possible and to obtain an estimate of the environmental variance from within each one. A snag in this context is that inbred lines have often been found to

be more variable than one would expect, on the basis of their supposed genetic uniformity. The problem is discussed by Biggers and Claringbold (1954) and by Grüneberg (1954). The reason for the extraordinary variability of many inbred strains is probably related to their unusual degree of homozygosity, which, according to Lerner's (1954) postulate of genetic homeostasis, renders them more sensitive to environmental sources of variation. By the same token, F_1s may be too insensitive, again relative to outbred stocks.

Any attempt to estimate V_E, in practice, should therefore not only utilize as many inbred lines as practicable but also the crosses between them. If the two provide different estimates of V_E, they should probably be averaged, although in some cases it may be preferable to use the F_1s only. It is impossible to provide objective guidance on this point; the decision must be a subjective one, based on the experimenter's knowledge of his material. The estimate of the environmental variance so obtained refers only to the population that the inbreds and their crosses represent. Its extrapolation to other populations may lead to wrong conclusions. But within such a population the component within strains or crosses estimates V_E, while the component between genotypes estimates V_G.

This partitioning of the total variance indicates the importance of the genotype in determining the phenotype. The ratio V_G/V_P indicates *the degree of genetic determination.* This ratio is sometimes referred to as "heritability in the broad sense," which is clumsy usage. Some writers have referred to it simply as "heritability," without a qualification, which is confusing. The term "heritability," as used in the genetic literature and throughout the remainder of this chapter, will be defined in a narrower, but more useful, sense below.

In the same way as G was split into $A + D$, so can V_G be partitioned:

$$V_G = V_A + V_D$$

It can be shown from Table 11.3 that A and D are not correlated, so that there is no covariance term. The term V_A is the variance of breeding values, associated with variation in the average (additive) effect of the genes. It is thus termed the *additive variance.* Its importance in the theory and practice of quantitative genetics is paramount. It is furthermore a useful concept, as it can be estimated directly in a population, in a way that will be described shortly. The ratio V_A/V_P is termed the *heritability*, which is therefore defined as the proportion of the phenotypic variance due to additive genetic sources. It is the main cause of resemblance between relatives, and it also indicates the reliability of the phenotype as a guide to breeding value.

Although V_A is termed the additive variance, it is worth stressing again that this does not imply additive gene action. It measures the variation in breeding value, which we have seen to depend upon, among other things, dominance.

The remainder of the genotypic variance, with respect to a single locus, is the variance of the dominance deviations. The algebraic ex-

pressions for the additive and dominance variances are obtained as follows:

"Variance" is defined as the mean-squared deviation from the population mean. As we have already determined both A and D in terms of deviations, in Table 11.3, all that is required is to square the values obtained, multiply by the frequency, and summate over the three genotypes.

Thus,

$$\begin{aligned} V_A &= 4p^2q^2\alpha^2 + 2pq(q-p)^2\alpha^2 + 4p^2q^2\alpha^2 \\ &= 2pq\alpha^2(p+q)^2 \\ &= 2pq\alpha^2 \\ &= 2pq[a + d(q-p)]^2 \\ V_D &= 4p^2q^4d^2 + 8p^3q^3d^2 + 4p^4q^2d^2 \\ &= 4p^2q^2d^2(q^2 + 2pq + p^2) \\ &= (2pqd)^2 \end{aligned}$$

What, then, do these expressions tell us about the additive and dominance variances? Firstly, let us examine their relative magnitudes. Since

$$V_A = 2pq[a + d(q-p)]^2$$

and

$$V_D = 2pq(2pqd^2)$$

it is seen that $V_D > V_A$ only when

$$2pqd^2 > [a + d(q-p)]^2$$

Substitution of values here will show that this will hardly ever occur except at intermediate gene frequencies when $d > a$, that is, in overdominant situations. The first conclusion is therefore that the additive effect of the genes usually contributes more to the variance than the dominance deviations. This conclusion is depicted graphically in Figure 11.3. Both expressions reveal another important conclusion, namely, that genes contribute more variance at intermediate than at extreme frequencies.

Extending the treatment to more than one locus, there is the additional complication of variance due to interaction effects. The theoretical consequences of interaction variance have not been developed very extensively, but the general conclusion seems to be that its contribution to genotypic variance is not very great. Interaction may occur between two loci or between more, but the more loci concerned, the less their relative contribution to the interaction variance. Although the partitioning is purely a theoretical one, several types of interaction variances are recognized, as this helps in the formulation of some problems. The interaction may be between the breeding values of two loci (V_{AA}), between the breeding value of one and the dominance deviations of the other (V_{AD}), between two dominance deviations (V_{DD}), and so forth, to multilocular situations (V_{AAA}, V_{AAD}, etc.). The full formula for the partitioning of phenotypic variance then becomes

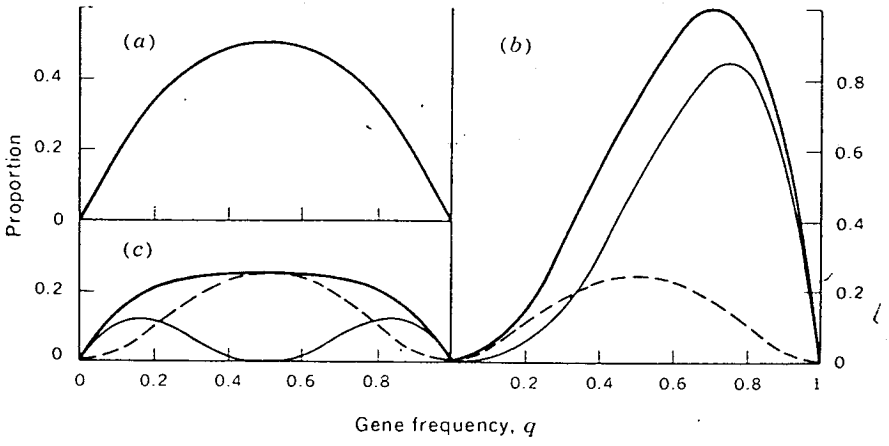


Figure 11.3 Magnitude of the genetic components of variance arising from a single locus with two alleles, in relation to the gene frequency. Genotypic variance, thick lines; additive variance, thin lines; dominance variance, broken lines. The gene frequency q is that of the recessive allele. The degrees of dominance are in (*a*) no dominance ($d = 0$); in (*b*) complete dominance ($d = a$); and in (*c*) "pure" overdominance ($a = 0$). The figures on the vertical scale, showing the amount of variance, are to be multiplied by a^2 in graphs (*a*) and (*b*), and by d^2 in graph (*c*). (*From Falconer, 1960, by courtesy of author and publisher.*)

$$V_P = V_A + V_D + V_I + V_E$$
$$= V_A + V_D + V_{AA} + V_{AD} + V_{DD} + \cdots + V_E$$

In practice, the dominance and interaction variances are grouped together into *nonadditive* genetic variance. As mentioned earlier, the really important division is usually into additive genetic variance and the remainder, nonadditive and environmental together. Occasionally, it may be possible to obtain an independent estimate of the environmental variance, in which case a realistic partitioning may be

$$V_P = V_A + V_{NA} + V_E$$

where V_{NA} is the nonadditive genetic variance. Beyond this we can seldom go. The additive variance is of prime importance. The nonadditive part sometimes contributes to the resemblance between relatives, as will be shown shortly, and it is also very pertinent to the study of inbreeding, to be discussed later.

Before we leave the subject of variance components, a word must be said about *environmental variance*. By definition, it refers to all variation not attributable to genetic causes, and the sources of it may be many. Where possible, genetic experiments should be designed to reduce the environmental variance, where likely sources of such variance are known. Nutrition, climatic, and housing factors are obvious examples; in mammals, maternal effects constitute another source of environmental variance. But over and above all the identifiable sources, there is usually still some residual environmental variance,

often a considerable amount. It is convenient in this context to recognize two kinds of environmental variance; one refers to factors that cause individuals to differ, the general *environmental variance* (V_{Eg}), while the other refers to variation within an individual, the *special environmental variance* (V_{Es}), for instance, variation from day to day or even from minute to minute. The latter is the reason for replicating measurements in order to assess the phenotypic value of an individual. Obviously, repeated measurements will reduce the V_{Es} component of V_P. If we now rewrite our formula as

$$V_P = V_G + V_{Eg} + V_{Es}$$

the variance of the mean of n repeated measurements becomes

$$V_{P(n)} = V_G + V_{Eg} + 1/n\, V_{Es}$$

We should perhaps reflect momentarily on the biological validity of repeated measurements. Behavioral scientists are well aware of the difficulties here. The implicit assumption is that the second and subsequent measurements do, in fact, measure the same effect as the first measurement. But it is possible that the very nature of the first measurement may render this assumption untenable.

These, then, are the sources of variance recognized either in practice or in theory. Variance components are the building blocks of quantitative genetics. Before discussing some aspects of methodology in quantitative genetics, we must first examine in more detail the causes of resemblance between relatives.

THE CAUSES OF RESEMBLANCE BETWEEN RELATIVES

In general, resemblance between groups, be they relatives or otherwise, is dependent upon the covariance of members of a group. This factor can be described in a variety of ways, the choice being usually one of convenience. For instance, where the groups concerned are groups of two, a product-moment correlation or a regression coefficient is the simplest expression to employ. But where groups consist of more than two individuals, then components of variance are extracted from an analysis of variance; they can be conveniently expressed as an intraclass correlation (t).

$$t = \frac{\sigma_B^2}{\sigma_W^2 + \sigma_B^2}$$

where σ_B^2 is the component of variance between groups and σ_W^2 is the component within groups.[2] To say that groups differ is another way of saying that members of a group are alike. The greater the

[2] Some readers may be unfamiliar with components of variance and intraclass correlations. An exposition of them may be found in many statistical textbooks, and a monograph by Haggard (1958) deals exclusively with the subject. It should be noted also that the symbol t for the intraclass correlation, used extensively in the genetic literature, is not to be confused with Student's t.

component between groups, the greater the relative similarity of members within a group. Thus the component of variance between groups can equally well be regarded, qualitatively at least, as the covariance of members of a group.

Just as variance was split up into genotypic and environmental components, so now must covariance be treated in the same way. We shall begin again with a single locus and later introduce the complication of interaction when we summate over loci. Two examples will be worked out, to illustrate the two ways of looking at covariance mentioned above. We shall examine first the genetic covariance, i.e., how the genotype of an individual is related to the genotype of a specified relative. Environmental covariance between relatives should be avoided by experimental design but, as we shall see shortly, this is not always possible.

Let us determine first the genetic covariance of an offspring with *one* of its parents, i.e., the covariance of their genotypic values. The genotypic value of, say, genotype A_1A_1, expressed as a deviation from the population mean, was given in Table 11.3 as $2q(a - pd)$. The algebra is simplified if we now express this as $2q(\alpha - qd)$ by substituting $a = \alpha - d(q - p)$. The mean genotypic value of the offspring of A_1A_1 is one-half of the breeding value of that genotype, i.e., ½ of $2q\alpha$. By extending this reasoning to the other two genotypes, we arrive at Table 11.4.

With the genotypic values now expressed as deviations, all that remains is to multiply the two together and also by the frequency, to obtain the covariance of an offspring with one of its parents (cov_{OP}). The expression simplifies to

$$\begin{aligned} cov_{OP} &= pq\alpha^2(p + q)^2 + 2p^2q^2\alpha d(-q+q - p+p) \\ &= pq\alpha^2 \end{aligned}$$

Since V_A was seen earlier to be $2pq\alpha^2$,

$$cov_{OP} = \tfrac{1}{2}V_A$$

This is intuitively reasonable; a parent's genetic contribution to an offspring is one-half the average effect of its genes. Over the whole population, therefore, the covariance of offspring with one parent is one-half the additive variance.

The other kind of relationship that we shall examine concerns the genetic covariance of half-sib groups, i.e., groups of progeny having one parent in common, usually the sire. The covariance among members of a half-sib group is, as explained earlier, the variance of the

Table 11.4

Genotype	Frequency	Genotypic value of parents	Genotypic-value of offspring
A_1A_1	p^2	$2q(a - qd)$	qa
A_1A_2	$2pq$	$(q-p)a + 2pqd$	$\frac{1}{2}(q-p)a$
A_2A_2	q^2	$-2p(a + pd)$	$-pa$

means of such groups. Given a parent of specified genotype, the mean value of its progeny, expressed as a deviation, is shown in Table 11.4. Squaring the deviations and multiplying by the frequency therefore give the variance of means of half-sib groups, this being the covariance (cov_{HS}) that we require:

$$\begin{aligned} cov_{HS} &= p^2q^2\alpha^2 + \tfrac{1}{2}pq(q-p)^2\alpha^2 + q^2p^2\alpha^2 \\ &= pq\alpha^2[\tfrac{1}{2}(p+q)^2] \\ &= \tfrac{1}{2}pq\alpha^2 \end{aligned}$$

Since $V_A = 2pq\alpha^2$,

$$cov_{HS} = \tfrac{1}{4}V_A$$

This again is intuitively reasonable, since two half sibs have, on an average, one-quarter of their genes in common. The additive effect of these genes is represented by $\frac{1}{4}\ V_A$.

By exactly the same principles, although a trifle more laboriously at times, genetic covariances among other types of relationships can be derived. For instance, the covariance between an offspring's value and the mean value of its two parents ($cov_{O\bar{P}}$) can be shown to be, again, $\frac{1}{2}\ V_A$. Although this covariance, in absolute terms, is the same as that with one parent, it is now of greater relative importance if it is expressed as a correlation or regression coefficient, since the variance of midparental values is only one-half the variance of the parents considered singly.

Another covariance of some importance is that among full-sib groups (cov_{FS}), which can be shown to be $\frac{1}{2}\ V_A + \frac{1}{4}\ V_D$, an example where the nonadditive genetic variance contributes to the resemblance between relatives. This is because full sibs share not only the additive effects of half their genes, accounting for the $\frac{1}{2}\ V_A$ term, but also a quarter of the dominance deviations, as a result of being identical for one-quarter of all the loci segregating in that mating.

Extending the treatment now to multilocular situations, the expressions for covariance must be expanded to accommodate the interaction terms. The reason for partitioning the variance into V_{AA}, V_{AD}, V_{DD}, etc., will now become apparent. Qualitatively, we may approach the question as follows: If the breeding values of two genes interact, giving rise to V_{AA}, the coefficient of V_{AA} in the expression for covariance is the probability of those two genes being present in the two relatives. For instance, if the genotype A_1–B_1– shows an interaction deviation, then in the covariance of an offspring with one parent, the probability of offspring and parent both containing A_1 or both containing B_1 is $\frac{1}{2}$, giving rise to the term $\frac{1}{2}\ V_A$, as before. The probability of both containing A_1 and also B_1 is $\frac{1}{4}$, giving a term of $\frac{1}{4}\ V_{AA}$. Thus,

$$cov_{OP} = \tfrac{1}{2}V_A + \tfrac{1}{4}V_{AA}$$

Similarly,

$$cov_{HS} = \tfrac{1}{4}V_A + \tfrac{1}{16}V_{AA}$$
$$cov_{FS} = \tfrac{1}{2}V_A + \tfrac{1}{4}V_D + \tfrac{1}{4}V_{AA} + \tfrac{1}{8}V_{AD} + \tfrac{1}{16}V_{DD} + \cdots$$

In general, if, in the expression for covariance, V_A takes the coefficient x, and V_D takes y, then the full formula to allow for interaction becomes

$$cov = xV_A + yV_D + x^2V_{AA} + xyV_{AD} + y^2V_{DD} + x^3V_{AAA} + \cdots$$

The general conclusion is that as V_I is usually rather small compared with V_A or even V_D, the interaction components do not figure prominently as a cause of resemblance between relatives. But in some situations, their effects may not always be negligible.

The last cause of resemblance between relatives is any similarity of the environment to which they may be subjected. As we depend upon the degree of resemblance between relatives to arrive at genetic conclusions, it is obvious that environmental causes of similarity must be excluded or else accommodated. Experimental design can often be employed to reduce the environmental covariance but, for mammals especially, total exclusion cannot always be ensured.

The component of environmental variance previously symbolized V_{Es} will not, by definition, contribute to the similarity of relatives; it refers to variation within an individual with respect to repeated measurements. But the general environmental variance V_{Eg} may cause trouble. It is convenient in this context to repartition the environmental variance into a component called *common environment* (V_{Ec}), tending to make members of a family more alike and therefore different families more different, and a second component, symbolized V_{Ew}, the *within-group component*, which causes individuals to differ irrespective of whether they are related or not. To avoid confusion, it should be noted that V_{Ew} contains all of V_{Es} and some of V_{Eg}, on the previous partitioning. We are concerned here only with V_{Ec}, the environmental influences common to members of a family or group. Possible sources of V_{Ec} come through housing, e.g., if related Drosophila are stored in the same bottle or mice in the same cage. In wild populations, relatives may tend to occupy the same habitat. Sources such as these may be allowed for, if recognized, and under laboratory conditions the component of variance due to common environment can often be assessed directly by setting up replicate bottles of flies, etc. In mammals, however, one cause of V_{Ec} that little can be done about comes through maternal effects, where full sibs, in particular, resemble each other not only because of their genetic covariance but also because the sibs occupied the same uterus in the foetal stage and suckled the same dam postnatally. Where maternal effects exist, the expression for the covariance of full sibs should be rewritten

$$cov_{FS} = \tfrac{1}{2}V_A + \tfrac{1}{4}V_D + \tfrac{1}{4}V_{AA} + \cdots + V_{Ec}$$

It is largely because of the common-environment component that, in genetic studies, full sibs tend to be of less use than the more distantly related half sibs. Full sibs have been singled out for comment, but other relationships should be closely examined in every experiment for possible sources of V_{Ec}, such as maternal effects or managemental factors. The common environment component may produce effects that

deserve study in their own right; maternal effects, for instance, as revealed by differences between reciprocal crosses, are often of considerable interest. But in genetic work, the common environment factor is usually best excluded where possible, for it tends to be confounded with genetic components in such a way that it may not be separable.

The covariance may be estimated from the measurable parameters of the population. Consider, firstly, the covariance of offspring with one parent. The variance of the parents is, by definition, the phenotypic variance of the population. Knowing V_P, then, we could assess cov_{OP} directly from the expression

$$b_{OP} = \frac{cov_{OP}}{V_P}$$

Similarly, the total variance in a population of sib groups is likewise V_P, and the component between groups has been shown to be the covariance of members of a group. The covariance of half sibs and of full sibs can therefore be estimated directly from the intraclass correlation (t). Lastly, the variance of midparental values ($V_{\bar{P}}$) is

$$\begin{aligned} V_{\bar{P}} &= V(\tfrac{1}{2}X + \tfrac{1}{2}Y) \\ &= \tfrac{1}{4}V_X + \tfrac{1}{4}V_Y \end{aligned}$$

where X and Y are the phenotypic values of the two parents. If

$$V_X = V_Y = V_P$$

then

$$V_{\bar{P}} = \tfrac{1}{2}V_P$$

Thus the covariance of offspring with the midparental value is obtained:

$$b_{O\bar{P}} = \frac{cov_{O\bar{P}}}{\tfrac{1}{2}V_P}$$

In practice, however, such covariances are not usually estimated, any more than covariances are normally assessed directly in statistical work. Genetic covariances, in absolute terms, are of little direct interest; their magnitudes vary, as we have seen, according to the closeness of the relationship. They are therefore reduced to a common base;

Table 11.5

Relationship	Covariance	Regression (b) or intraclass correlation (t)
Offspring: one parent	$\tfrac{1}{2}V_A$	$b = \frac{\frac{1}{2}V_A}{V_P} = \tfrac{1}{2}h^2$
Offspring: midparent	$\tfrac{1}{2}V_A$	$b = \frac{\frac{1}{2}V_A}{\frac{1}{2}V_P} = h^2$
Half sibs	$\tfrac{1}{4}V_A$	$t = \frac{\frac{1}{4}V_A}{V_P} = \tfrac{1}{4}h^2$
Full sibs	$\tfrac{1}{2}V_A + \tfrac{1}{4}V_D + \cdots + V_{Ee}$	$t = \frac{\frac{1}{2}V_A + \frac{1}{4}V_D + \cdots + V_{Ee}}{V_P} > \tfrac{1}{2}h^2$

since V_A figures prominently in them all, the practice usually is to express V_A as a fraction of V_P. This ratio was described earlier as the heritability. Table 11.5 shows how the covariance is related to the heritability (symbolized h^2).

This, then, shows in principle how estimates of heritability are derived. It shows also the final product of our examination of variances and covariances, and how different relationships can be analyzed and unified into one ratio that refers to the whole population. Before we examine in more detail how heritabilities are determined in practice, we should examine the importance and usefulness of heritability as a concept and its role in quantitative methodology.

HERITABILITY

We have just seen how the formulation of the resemblance between relatives leans heavily on the heritability, as the additive genetic variance is the main cause of the resemblance. The greater the heritability, the greater the covariance and therefore the similarity between relatives. Secondly, the magnitude of the heritability, ranging from 0 to 1, determines the reliability of the phenotype as a guide to breeding value, something which confers a predictive role on the heritability. To see this, let us determine the regression of breeding value on phenotype. Let

$$P = A + R$$

where R is the remainder, nonadditive genetic and environmental. Then

$$\begin{aligned} cov_{AP} &= cov_{A(A+R)} \\ &= cov_{AA} = V_A \end{aligned}$$

since $cov_{AR} = 0$.

So

$$b_{AP} = \frac{cov_{AP}}{V_P} = \frac{V_A}{V_P} = h^2$$

The higher the heritability, the more accurate is the phenotype as a guide to breeding value. Given the phenotypic value, the breeding value can be predicted from the heritability according to the formula

$$Exp\ A = b_{AP}P = h^2P$$

This leads directly to another predictive role, that of response to selection. Expressing both A and P now as deviations, if S (selection differential) is the deviation of the parents from the population mean, the expected deviation of the progeny, R (response to selection), is given by

$$Exp\ R = h^2S$$

since the breeding value is, by definition, the mean genotypic value of the offspring.

The concept of heritability has, therefore, two connotations, one descriptive and one predictive. The first describes the proportion of the total phenotypic variance that is additive genetic, and it leads to a causal description of the resemblance between relatives. The second indicates the reliability of the phenotype as a guide to breeding value, and it leads to the prediction of the results of certain manipulative processes. Quantitative genetics in one of its most pragmatic aspects, namely, animal breeding, makes extensive use of the concept of heritability in the development of its theory. But it is no less important in a more "fundamental" aspect, the investigation of the genetics of a quantitative character. Without the assessment, in some form, of the additive genetic variance—the variance due to the average effect of the genes in the population—the genetic study of such characters can scarcely begin.

In the context of heritability, it is important to remember the basic raw materials that went into the process, namely, a, d, q, and V_E. This means that any estimate of heritability refers to one population only and will not necessarily hold if, for instance, the gene frequencies are changed. The extent to which the results from one population can be extrapolated to another depends on the similarity, both genetic and environmental, of the two populations. Nevertheless, as far as laboratory animals and also domestic livestock are concerned, there is generally wide agreement on the relative magnitude of the heritability for a given character in a given organism, with the following general conclusion. Characters that have been subjected to natural selection, i.e., components of natural fitness, such as fertility or maternal performance, tend to have low heritabilities. The suggestion is that the additive variance has been largely exhausted through the action of natural selection in the past. On the other hand, characters such as the fat content of milk in cattle or bristle number in Drosophila, which do not seem to be such direct components of natural fitness, have in general much higher heritabilities. It will be interesting to see, as more information accumulates, how some behavioral characters fit into the picture.

THE ESTIMATION OF HERITABILITIES

Let us examine a little more closely how heritabilities are estimated in practice, although the principles on which the methods rest have already been given.

The first step involves the practical decision of the type of relationship to employ. The overriding concern at this stage is to avoid environmental sources of covariance that would lead to the wrong answer by inflating the estimate of the heritability. If maternal effects, for instance, are known to affect the character, the use of offspring and dams or of full sibs is vitiated, and other types of relationships, such as sire and offspring or half sibs, must be chosen for the estimation. Such procedure can be indicated only by experience of the biological nature of the character. Other sources of environmental

covariance, which may result from feeding, housing, handling, etc., should be excluded as far as possible. Their effect will differ for different characters, and there is no substitute for common sense in avoiding the pitfalls in this respect.

The first method of heritability estimation implicit in Table 11.5 is the regression of offspring on parent. Whatever the type of relationship involved, it is of course bound to include parents and offspring and with the possible biases mentioned above in mind, and provided that the parents have been measured at the right age, etc., it is usually worthwhile to obtain one estimate of the heritability by the regression method, even though the structure has been designed for a sib analysis or some other type of relationship. Little more need be said about the regression method, as the regression coefficient of the mean value of the offspring on the value of one parent measures one-half the heritability, while the regression of offspring on the mean value of the two parents measures the heritability itself. The regressions are calculated from paired observations in the usual way. However, three possible modifications of the regression method should be mentioned.

The first modification is applicable if the mean values of the offspring are based on families differing in size. These can be weighted, so that rather more attention is given to the larger families, according to methods suggested by Kempthorne and Tandon (1953) or by Reeve (1955). Usually such adjustments do not greatly alter the estimate of the heritability, but they do increase the precision of the estimate if the families differ widely in size.

The second modification concerns the variances of the two sexes. It was assumed in our treatment of covariance that these were equal. If they are not, two separate regressions should be taken, that of sons on sires and that of daughters on dams. The average of the two will then provide an overall estimate of heritability, though this will be rather a meaningless figure unless the two do not differ.

The third possible modification becomes imperative if a sire is mated to more than one dam, as would be the case, for instance, in a half-sib structure. The mean of the offspring of one mating could not then be regressed on the midparental value, as one of the parents would be common to several paired observations. The mean of all the sire's offspring could still be regressed on the sire's own value, but the number of sires is seldom sufficient to make this profitable. Under the circumstances, the procedure is to regress the offspring on the dam within each sire group, pooling the degrees of freedom and the sums of squares and of products to obtain a weighted average regression. Such a regression, involving basically only one parent, estimates half the heritability, as before, and is known as the *intrasire regression of offspring on dam.*

The regression method of estimating heritability is therefore conceptually straightforward. We shall now examine the other method of heritability estimation suggested by Table 11.5. This involves the derivation of the intraclass correlation from the analysis of variance of a sib structure, and its theory is somewhat more complicated.

Table 11.6

Source	Degrees of freedom	Composition of mean square
Between sires	$s-1$	$\sigma_W^2 + k\sigma_D^2 + dk\sigma_S^2$
Within sires		
Between dams	$s(d-1)$	$\sigma_W^2 + k\sigma_D^2$
Within dams	$sd(k-1)$	σ_W^2

In practice, the breeding structure usually involves a mixture of full sibs and half sibs. If s sires are each mated to d dams, each of which produce k offspring for measurement, the structure lends itself to a standard hierarchical analysis of variance, as shown in Table 11.6. If the ds and ks are unequal, then there are standard statistical techniques to adjust the coefficients of the components as they affect the mean squares. Note also that the symbol σ^2 is employed here to denote observational components of variance, to distinguish them from causal components, symbolized V.

From the analysis, we can now derive three components of variance, one between sires (σ_S^2), one between dams (σ_D^2), and a third one within groups of full sibs (σ_W^2). The next step is to relate these observational components to the causal components that we have derived previously and thereby derive the ratio V_A/V_P, the heritability.

Firstly, it follows by definition that the sum of the observational components estimates the total phenotypic variance σ_T^2, though the sum may not tally exactly with the total variance as observed.

$$\sigma_T^2 = \sigma_S^2 + \sigma_D^2 + \sigma_W^2 = V_P$$

Secondly, σ_S^2 measures, in fact, the component of variance of half-sib groups, since it is in the nature of the analysis that the deviations of full sibs from their own family means are removed separately. Thus, σ_S^2, the component between half sibs, is equal to the covariance of half sibs, which we saw earlier to be $\frac{1}{4}V_A$. Thirdly, σ_W^2, the component of variance within full-sib groups, is the complement of the component between full-sib groups, $\sigma_{B(FS)}^2$, and the two must add up to the phenotypic variance:

$$V_P = \sigma_W^2 + \sigma_{B(FS)}^2 = \sigma_W^2 + cov_{FS}$$

Thus

$$\begin{aligned}\sigma_W^2 &= V_P - cov_{FS} \\ &= V_P - \tfrac{1}{2}V_A - \tfrac{1}{4}V_D - V_{Ec} \\ &= \tfrac{1}{2}V_A + \tfrac{3}{4}V_D + V_{Ew}\end{aligned}$$

Table 11.7

Observational component	Causal component
$\sigma_S^2 = cov_{HS}$	$\frac{1}{4}V_A$
$\sigma_D^2 = cov_{FS} - cov_{HS}$	$\frac{1}{4}V_A + \frac{1}{4}V_D + V_{Ec}$
$\sigma_W^2 = V_P - cov_{FS}$	$\frac{1}{2}V_A + \frac{3}{4}V_D + V_{Ew}$
$\sigma_T^2 = V_P$	$V_A + V_D + V_{Ec} + V_{Ew} = V_P$

Lastly, σ_D^2 may be obtained by subtraction:

$$\begin{aligned}\sigma_D^2 &= V_P - \sigma_W^2 - \sigma_S^2 \\ &= V_P - (V_P - cov_{FS}) - cov_{HS} \\ &= cov_{FS} - cov_{HS}\end{aligned}$$

These derivations are summarized in Table 11.7.

The interaction variance could easily be included among the causal components by reference to the genetic covariances given previously; it has been neglected here to simplify the presentation, and its effect, in any case, is not usually very great.

We thus see how the total variance can be partitioned into observational components which in turn can be equated to causal components. The heritability can now be obtained from the analysis. Since

$$\frac{\sigma_S^2}{\sigma_T^2} = \tfrac{1}{4}\frac{V_A}{V_P} = \tfrac{1}{4}h^2$$

then,

$$h^2 = \frac{V_A}{V_P} = \frac{4\sigma_S^2}{\sigma_T^2}$$

Unfortunately, the other causal components are not deduced so easily, since there are basically only three equations for four unknowns. σ_D^2 could be utilized in exactly the same way to obtain, in this case, an upper limit to the heritability. If, in fact, the estimate thereby obtained does not greatly exceed that obtained from the between-sire components, it indicates that neither V_D nor V_{Ec} is very important. But V_D, V_{Ec}, and V_{Ew} can be estimated from this analysis only if they are good reasons for believing that one of them can be safely neglected and equated to zero.

SOME CONSIDERATIONS OF EXPERIMENTAL DESIGN IN THE ESTIMATION OF HERITABILITIES

Having obtained an estimate of a heritability, one requires some indication of its reliability. This becomes a statistical question of attaching standard errors to regression coefficients or to intraclass correlations, as the case may be. Formulae for these standard errors are given in the statistical literature, and without going into detail, we must examine what they tell us about experimental design. Much of the material in this section is discussed in more detail by Robertson (1959).

Sheer physical considerations limit the size of any experiment, since the number of animals that can be measured is restricted either by space or by the time and labor involved in the measurement. However, the restrictions so imposed still leave room for the manipulation of family size, and a choice has to be made between measuring a few families accurately, i.e., by the use of a large number of offspring per family, or measuring more families less accurately. The two must be balanced to give the optimal design, which is the design that mini-

mizes the sampling variance of the heritability estimate. We shall deal with the regression and intraclass methods, in turn, as before. We shall employ the following symbols in this section:

Y = mean value of offspring
X = value of one parent or mean value of two parents
σ_Y^2 and σ_X^2 = respective variances
n = number of offspring per parent
N = number of families
T = total number measured

Thus, if n offspring and one parent are measured per family, $T = N(n + 1)$; if both parents are measured, $T = N(n + 2)$. In sib structures, where the parents need not be measured, $T = Nn$. It is supposed that T will be limited by the total facilities available, and different methods can therefore be compared for efficiency, given the same total facilities.

Let us consider first the regression technique, as utilized to determine the heritability from the relationships:

$b_{YX} = \frac{1}{2}h^2$ (for one parent) or h^2 (for midparent)

Let the sampling variance of b be denoted by σ_b^2. It is well known that

$$\sigma_b^2 = \frac{1}{N-2}\left(\frac{\sigma_Y^2}{\sigma_X^2} - b^2\right)$$

The subsequent algebra is simplified if an approximation is derived on the basis that N is fairly large and the b^2 is fairly small, as both will tend to be. Thus,

$$\sigma_b^2 \approx \frac{1}{N}\frac{\sigma_Y^2}{\sigma_X^2}$$

Now, σ_X^2 is V_P in the case of one parent, and $\frac{1}{2}V_P$ for the midparental values. But σ_Y^2, the variance of the mean of offspring groups, depends on the number in the group and the intraclass correlation (t) between members of the group. It can be shown to be

$$\sigma_Y^2 = \frac{1 + (n-1)t}{n} V_P$$

(That this formula is reasonable can be appreciated by substituting $n = 1$ or $t = 0$. The values for σ_Y^2 then become V_P or V_P/n, as expected.) We can now substitute in the formula for σ_b^2:

$$\sigma_b^2 \approx \frac{1 + (n-1)t}{nN}$$

for the regression on one parent. It is twice as great for the regression on midparent. Assuming T to be fixed, it can be shown that σ_b^2 is minimal when $n = \sqrt{(1-t)/t}$ for one parent, or $\sqrt{2(1-t)/t}$ for midparent.

This, then, gives the optimal number of progeny to measure per family and is now seen to depend on t. The intraclass correlation is of course closely related to the heritability, which is not known when the experiment is designed. The optimal design can therefore be derived only a posteriori, but limits to the optimal value of n can be obtained a priori, since t for full-sib families can vary only between 0 and $\frac{1}{2}$, in the absence of complications. Substitution of possible values of t will show that, if the offspring are regressed on one parent, about 10 offspring per family should be measured if the heritability is around 2 percent, while the number drops to 2 or so for a heritability around 50 percent. If midparental values are employed, the numbers are 14 and, still, about 2, respectively. In complete ignorance of what to expect, a number of 5 or so is a reasonable one to employ in order to minimize the sampling variance.

The optimum structure that we have derived is seen to be independent of the total facilities available. We can now derive the standard error of the heritability estimate, in terms of T. For illustration, we shall consider a character of 20 percent heritability, so that $t = 0.1$. The optimum n then is 3 for single-parent regression, and 4 for midparent. Bearing in mind that, since $h^2 = 2b$, $\sigma_{h^2}^2 = 4\sigma_b^2$, the formulas already given can be manipulated to show that

$$\text{For single-parent regression: } \sigma_{h^2}^2 = \frac{6.4}{T}$$

$$\text{For midparent regression: } \sigma_{h^2}^2 = \frac{3.9}{T}$$

Two intuitively obvious conclusions emerge: First, the larger the value of T, the smaller the sampling variance, and, second, estimates derived from midparental values are more precise, for given total facilities.

The final point concerning the regression method is what precision is required to make the experiment worthwhile. Continuing with our example of a heritability of 20 percent, suppose we require a standard error of not greater than 10 percent. This is not very ambitious—just sufficient to demonstrate that the heritability is not zero. Then, if σ_{h^2} is 0.1, $\sigma_{h^2}^2$ becomes 0.01. We can now determine T, and since $T = N(n + 1)$ or $N(n + 2)$, where n is 3 or 4, respectively, we can calculate N to be 160 for single-parent regressions, and 65 for midparent regressions.

These results are summarized in Table 11.8, which shows the number of animals required to estimate a heritability of 20 percent with a standard error of 10 percent.

Table 11.8

	Single parent	Midparent
Number of parents measured per family	1	2
Number of offspring measured per family (n)	3	4
Number of families required (N)	160	65
Total number of animals measured (T)	640	390

These, then, are the minimal requirements, in terms of animals, in order to estimate a heritability of 20 percent within the broadest acceptable limits. Any increase in the precision of the estimate would require more facilities.

Turning now to sib analyses, utilizing the intraclass correlation for estimating the heritability, we can deal with two simplified structures within the one framework. These are either half-sib or full-sib families but not a mixture of both; this last case will be mentioned briefly later. A simple half-sib structure involves mating a sire to several dams, each of which provides one offspring for measurement. In a full-sib structure, each sire is mated to one dam only, and the mating provides several offspring. The number of families (N) is then equivalent to the number of sires in each case. If each family consists of n offspring, the total number of animals measured (T) is Nn, as the parents for these analyses do not need to be measured. From the analysis of variances between sires, we derive the intraclass correlation. The use of full sibs implicitly assumes that dominance and common environment are unimportant. Granted this assumption, then, the intraclass correlation in the case of full sibs estimates $\frac{1}{2}h^2$. In the case of half sibs, $t = \frac{1}{4}h^2$. Now the variance of the intraclass correlation (σ_t^2) is found in the statistical literature to be

$$\sigma_t^2 = \frac{2[1 + (n-1)t]^2 (1-t)^2}{n(n-1)(N-1)}$$

If both N and n are fairly large, as they will tend to be, this formula can be approximated without much loss of accuracy to

$$\sigma_t^2 \approx \frac{2(1+nt)^2 (1-t)^2}{nT}$$

Expressed in this way, σ_t^2 can be shown to be at a minimum when $nt = 1$, or $n = 1/t$.

As in the case of regression methods, the optimum structure is again seen to depend on the intraclass correlation. It is a small step now to express n, the optimum number of offspring to measure per family, in terms of the heritability; in the case of full-sib families, the optimum n is $2/h^2$, and $4/h^2$ in the case of half sibs. Thus, for a heritability of 20 percent, 10 offspring per family should be measured if full sibs are employed, and 20 offspring where half sibs must be used. But as the heritability is unknown at this stage, the optimum structure can be achieved only fortuitously, as before. In the complete absence of any knowledge of what to expect, experiments should be based on 10 to 15 full sibs or 20 to 30 half sibs, as the case may be. This, on average, seems to lead to the least loss of information. It should be recognized, however, that these optimal structures are not always easy to attain, and the size of sib groups is often dictated more by the reproductive capacity of the organisms than by statistical considerations.

The sampling variance of the heritability estimate can be deduced as follows: Given the optimum structure of $nt = 1$, the above formula for σ_t^2 can be simplified further:

$$\sigma_t^2 \approx \frac{8(1-t)^2}{nT} \approx \frac{8}{nT} = \frac{8t}{T}$$

since we can neglect $(t^2 - 2t)$ with little loss of accuracy.

This leads directly to the sampling variance of the heritability estimate:

For full sibs: $\sigma_{h^2}^2 = 4\sigma_t^2 = \frac{16h^2}{T}$

For half sibs: $\sigma_{h^2}^2 = 16\sigma_t^2 = \frac{32h^2}{T}$

We can now determine the scale of experimentation required to achieve a standard error of a given magnitude. Continuing with our previous example of a heritability of 2 percent and presuming that we are aiming, as before, at the modest objective of reducing σ_{h^2} to 0.10, or $\sigma_{h^2}^2$ to 0.01, we can substitute this value in the formulae and solve for T. From the relationship $T = nN$, we can further calculate N, the number of families that should be measured. The results are summarized in Table 11.9.

Full sibs, where they can be employed, are therefore twice as efficient as half sibs. It will be noticed also, referring back to Table 11.8, that, for the specific example chosen, a half-sib structure is of the same efficiency as the regression on one parent. A choice has often to be made between these two methods, and by substituting values in the formulae we have derived, it can be established that the following general rule applies. For a given total number of animals measured, a half-sib structure with optimal design gives a more accurate estimate of the heritability than the regression of offspring on one parent if the heritability is below 20 percent; for higher heritabilities, the opposite holds.

The half-sib structure just discussed, namely, one offspring per dam, becomes the most efficient one to use if only the component of variance between sires can be employed to estimate the heritability. If it is desired to use the between-dam component as well, the situation becomes more complicated, as the family group will then consist of a mixture of half and full sibs. Under these conditions, it can be shown that the optimal design is to mate three or four dams per sire, with $2/h^2$ offspring measured per dam. In the absence of any prior estimate of the heritability, about 10 offspring per dam should be measured.

Table 11.9

	Full sibs	Half sibs
Number of offspring measured per family (n)	10	20
Number of families required (N)	32	32
Total number of animals measured (T)	320	640

This sketchy consideration of experimental design, with rather crude algebra at times, leads to one unambiguous conclusion. It is that estimates of heritability become meaningless if they are based on small numbers. Even when the number of animals rises into the hundreds, the estimates are still not very precise. It may not always be possible to collect all the required data in one generation; in these circumstances, it may be expedient to pool information from more than one generation, on the assumption that the sources of variance do not alter in the meantime.

The implications of this section on experimental design, with specific regard to the estimation of heritabilities, can be summarized as follows:

1 Sheer physical considerations impose a limit on the total facilities available for any experiment. But it is possible to manipulate the variables, especially the breeding structure, to maximize the information that may be gained from these facilities. The concern is to reduce the sampling variance of the estimate of the heritability to a minimum.

2 Given the optimum structure for an experiment, it is possible to predict, within limits, the standard error of the estimate of the heritability. This gives the investigator a realistic idea of the scale of experimentation necessary, lest he should embark on a program doomed to futility from the start.

THE PROVISION OF MATERIAL FOR RESEARCH

So far in this chapter, we have been concerned with the genetics of quantitative characters in a static situation; we have examined ways of describing the genetics of a population as we find it, with respect to any character in which we may be interested. This descriptive approach is calculated to shed light on the inheritance of the character and to discover how the genetic variables affect the level of its expression. Paramount among the genetic variables are the gene frequencies. We must now examine ways in which gene frequencies can be manipulated to alter the level of expression of a character and thereby extend our understanding of the genetic control of particular biological compositions.

In the remainder of this chapter, we shall concern ourselves mostly with two agencies that can change the gene frequencies rather rapidly. The first is inbreeding, the effects of which are dispersive, resulting from random changes in gene frequencies within lines, though the overall frequencies in a population of such lines do not change. The second is selection, resulting in directional changes in gene frequencies, brought about by the differential fertility of individuals. We shall not consider mutation, as its effect on a quantitative character in the absence of selection is unimportant.

It should perhaps be said here that studies of inbreeding and of selection are not, on their own, potent genetic methodologies. The

additional conclusions that they permit about the genetics of a population are seldom rigorous, for the reason that the changes that they bring about defy close scrutiny. Similar results are often the products of different situations, and repeated experiments do not always provide the same answers. This is a reflection of the fact that sampling errors play a large part: first, in the determination of the genetic composition of the base population and, second, in any subsequent manipulation. Nevertheless, information from inbreeding and from selection programs is of value, and the results are of cumulative importance. But no doubt the chief use of inbreeding and of selection is in the provision of special strains for specific research requirements. The particular kind of research for which these strains are employed depends, of course, upon the organism and upon the character involved. For instance, the usefulness of strains of mice susceptible to cancer needs no elaboration; they have been employed in a multitude of ways in cancer research. Again, in their book, Fuller and Thompson (1960) refer several times, and in different contexts, to Tryon's maze-bright and maze-dull rats, illustrating the usefulness of this kind of material in research programs. Instances such as these could be multiplied to illustrate that the end products of inbreeding or selection are often of more value than any information about the route whereby the end product was obtained. Because of this, we shall deal with inbreeding and selection, in turn, from the point of view of developing special strains, as well as deriving information about the genetics of the population.

INBREEDING

Inbreeding can be defined as the union of gametes containing alleles identical by descent. Alleles are said to be identical by descent when they are the division products of one such allele that occurred in the past. The *coefficient of inbreeding* (symbolized F) is the probability that the uniting alleles are identical by descent and can be regarded as a correlation (ranging from 0 to 1) between the uniting gametes. This definition of inbreeding is, however, a theoretical one, and in practice it must be modified. For who knows whether or not any two alleles that are alike are identical by descent, if traced sufficiently far back? Therefore a base line must be fixed arbitrarily, and beyond this line no ancestries will be traced. Any specified degree of inbreeding then becomes one relative to this base population, which by definition has an inbreeding coefficient of zero. Likewise, the practical definition of inbreeding becomes the mating of individuals that are more closely related than the average relationship between all the individuals of that population.

The effects of inbreeding are widely known and do not require detailed comment here. Inbreeding is a dispersive process, in the sense that homozygotes are increased at the expense of heterozygotes. It tends to fix the alleles at a particular locus; i.e., they all become alike, and

further heterozygosity can arise only through mutation, which will initially occur in one individual only. The chance of fixation of any one allele is proportional to its initial gene frequency. Thus, if a number of lines from a base population are inbred simultaneously, a proportion p will become fixed for the A_1 allele, and q, or $1 - p$, for allele A_2, where p and q are the respective gene frequencies. If we consider a second locus, B_1B_2, with initial frequencies r and s, respectively, a proportion pr of lines will become fixed for A_1 and B_1, ps for A_1B_2, etc. As the number of loci increases, the probability of two lines being fixed for all the same alleles soon becomes negligibly small; this, perhaps, is the way in which the dispersive process is most clearly visualized.

Although the overall probability of fixation is determined by the initial gene frequency, it is purely a matter of chance which particular lines are fixed for A_1, which for B_2, etc. Because of this random dispersion, the final genetic composition of a line is unpredictable. It is presumably on this account that some inbreeding programs are subjected to simultaneous selection toward some given phenotype. In the light of present-day knowledge, however, this practice seems to have little to commend it; it is roundly condemned, with specific reference to behavior, by Broadhurst (1960). For one thing, selection is much more potent on its own when unaccompanied by the opposing dispersive effects of inbreeding. This is because a favorable gene may begin to become associated with an unfavorable one, and under a system of close inbreeding, they tend to remain so, despite selection. And even with only a few loci involved, this is a likely occurrence. For the establishment of strains for further research, inbreeding as such has little to contribute, unless homozygosity or random dispersion become ends in themselves, as indeed they are in special cases. And quite apart from all this, inbreeding has other ill effects, to be mentioned shortly.

Inbreeding, in the sense of increasing the homozygotes at the expense of heterozygotes, occurs in small populations, for it is intuitively obvious that small numbers make it more likely that uniting alleles are identical by descent. Theoretically, the change in inbreeding coefficient (ΔF) from one generation to the next can be shown to be, approximately,

$$\Delta F = \frac{1}{8N_M} + \frac{1}{8N_F}$$

where N_M and N_F are the numbers of male and female parents, respectively. Thus, five breeding pairs increase the inbreeding coefficient by 5 percent per generation, so that, with even so small a number, the increase in the inbreeding coefficient is not alarming. We must, however, distinguish here between rapid inbreeding, such as occurs when sibs or other close relatives are mated, and slow inbreeding through, say, a restriction of the population size. Selection, particularly natural selection favoring heterozygotes, has much more scope

when the inbreeding is slow. Thus, if heterozygotes, which in this context include heterozygous segments of chromosome, have any advantage in fitness, the ensuing natural selection will retard the approach to homozygosity. In fact, Hayman and Mather (1953) showed that only a moderate advantage of the heterozygotes will prevent complete fixation. It is also a matter of observation that small populations, e.g., stocks of laboratory mice, do not in fact suffer much from the effects of slow inbreeding, although they might be expected to have accumulated such effects through time, when the number of breeding pairs is only 10 to 20 per generation. This indicates that natural selection is, in fact, at work.

What, then, are the effects of inbreeding, with respect to the measurable parameters of the population? Starting again with a single-locus model, we shall examine first the effect of inbreeding on the population mean. When the inbreeding coefficient is raised from the arbitrary zero level to a value F, the relative frequencies of the three genotypes are modified, according to well-known formulae, in the direction of increasing the homozygotes at the expense of the heterozygote. If the initial frequency of the A_1 allele is again p, and that of the A_2 allele q, the modified frequencies when the inbreeding coefficient stands at F are shown in Table 11.10. It should be noted that these are average frequencies which refer either to one particular locus in a population of many inbred lines or, alternatively, to the array of loci, similar in kind and magnitude of effect, within any one line.

Multiplying the frequency by the value and summating over the genotypes, we can derive the population mean (M_F) when the inbreeding coefficient is F and also compare it with the mean (M_0) that we previously derived for zero inbreeding. For one locus,

$$m_F = a(p - q) + 2dpq - 2Fpqd$$
$$= m_0 - 2Fpqd$$

Generalizing by summating over loci,

$$M_F = M_0 - 2F \Sigma pqd$$

This shows that the effect of inbreeding to coefficient F is to reduce the mean, in terms of our arbitrarily assigned values, by $2Fpqd$. This formula permits three conclusions:

1 The change in the mean is linearly related to F; this is a theoretical conclusion, and it would perhaps be true to say that experimental support for it is not overabundant.

Table 11.10

Genotype	Frequency	Value
A_1A_1	$p^2 + Fpq$	$+a$
A_1A_2	$2pq - 2Fpq$	d
A_2A_2	$q^2 + Fpq$	$-a$

2 If d, or rather Σd, is zero, the mean does not change on inbreeding. Thus, genes that act additively, although redistributed between genotypes, do not affect the population mean on inbreeding.

3 The minus sign in the formula indicates that the change in the mean is always in the direction of the recessive allele. To the extent that recessive genes tend to be deleterious, the effect of inbreeding on the population is also deleterious.

There is a plethora of literature on the formal study of mutant genes, underlining the generally deleterious effects of recessives. This accords well with experimental experience of inbreeding, the deleterious effect of which is often all to obvious.

The decline in the population mean on inbreeding is known as *inbreeding depression*, the knowledge of which preceded its understanding of many centuries. It is particularly obvious in the case of reproductive capacity and maternal performance, the most obvious components of natural fitness. It was noted earlier that the additive genetic variance of such characters is relatively small but that their dominance variance is correspondingly greater. It is the source of this variance, the dominance deviations, that is the main cause of inbreeding depression. Furthermore, it is not sufficient that these deviations exist. They must also affect the character predominantly in the one direction; i.e., the dominant alleles must have a tendency to increase the phenotypic measurement and the recessive alleles to decrease it, or vice versa. The character is then said to exhibit *directional dominance*. It is intuitively acceptable that directional dominance should be characteristic of the components of natural fitness; alleles have themselves evolved, and the favorable ones are supposed to have evolved a dominant expression.

The establishment of directional dominance is about the only strict genetic conclusion that can be derived from the study of the effects of inbreeding on the population mean. If the mean does not change, then either all the genes act additively or the dominance deviations, on average, cancel each other out, that is, $\Sigma d = 0$. A third reason could apply if the inbreeding is sufficiently slow, namely, selection favoring the heterozygotes. In this case, the inbreeding coefficient, as calculated, would not reflect accurately the stage of gene dispersion in the population.

Turning now to the effects of inbreeding on the genotypic variance and its components, there is surprisingly little that may be said with profit in this context. The reason for this is twofold. Firstly, the theory has not yet been very fully developed. Secondly, as we have seen in the context of experimental design, variance components are very difficult to estimate with any precision, so that the experimental evidence on the subject is not very revealing. But in a general way, the tendency is for the genotypic variance of a population to be repartitioned on inbreeding, until ultimately it vanishes within lines and it all appears as the component of variance between lines. If all the genotypic variance is additive, which can happen only when all the genes affecting

a character act additively, then the expressions for the between- and within-line components of genotypic variance, for partial inbreeding, are as follows:

Between lines: $2FV_G$
Within lines: $(1 - F)V_G$
Total: $(1 + F)V_G$

Thus, when F reaches 1, the total genotypic variance is in fact twice what it was originally, on account of the increased number of homozygotes. But these expressions are true only when all the genotypic variance is additive. They do not hold for the additive effects of genes with dominance; they do not hold for the additive part of the genotypic variance if dominance variance also exists. The case of fully dominant genes, for instance, is quite different. Robertson (1952) showed that, in this case, the within-line variance rises until F is in the region of 0.5, as a result of the segregation of more homozygotes: it then declines. We need go no further to appreciate the futility of attempting to draw genetic conclusions from the study of variances during inbreeding, though empirical observations are always of interest if they are reasonably precise.

The subject of inbreeding should perhaps not be dismissed without a mention of its complement: *heterosis* or *hybrid vigor*. The two terms should be regarded as synonymous, for although some writers attach slightly different shades of meaning to them, there is no consistency among their practices. Heterosis, as a topic, rightly belongs in the provinces of plant and animal breeding, and it is difficult to see how its study, at this stage, can lead in any way to a deeper understanding of the genetics of a particular character. Suffice to say that heterosis among crosses can occur only for characters that display inbreeding depression, and its existence can therefore reflect no more than the presence of directional dominance.

It is sometimes mistakenly supposed that heterosis is somehow contingent upon the presence of overdominance, where the heterozygote exceeds in value either homozygote. This is not a prerequisite of heterosis; crosses between lines fixed for the recessive alleles at different loci are prone to exceed the level of either parental strain, given directional dominance. It is nevertheless true that overdominant loci, if they exist, may influence greatly the degree of directional dominance; a few loci overdominant in one direction may easily outweigh more loci that are dominant, or partially so, in the other. If and when it is present, overdominance may well have an overriding influence on the change in the mean during inbreeding and crossing. But the existence of overdominance in quantitative genetics is not easy to establish; it is difficult to distinguish from epistasis and impossible, in the short run, to distinguish from close repulsion linkage.

Heterosis, then, reflects directional dominance. It may occasionally be a means of providing research material if, for instance, uniform genotypes are required and inbred material proves to be too infertile

or too susceptible to sources of environmental variation. This exploitation of heterosis, however, can be approached only empirically, where possible practical advantages can be envisaged.

RESEARCH WORK WITH INBRED MATERIAL IN QUANTITATIVE GENETICS

The usefulness of isogenic material in many types of research needs no emphasis. In quantitative genetics, however, this usefulness is severely circumscribed by the limitation on the number of genotypes that are available. But even more important is the peculiar nature of inbred material and its derivatives.

The gametes of all the individuals of an inbred strain are all exact replicates of one another, except for the sexual dimorphism, which in this context is of negligible significance. Since no genes have been added to the population during inbreeding, it is possible that an exact replicate of this particular gamete could have been found in the base population of outbred individuals. In a sense, therefore, an inbred line can be considered as a representative of one gamete only from the base population, and an experiment involving, say, 10 lines is an experiment on a sample of 10 gametes out of a possible very large number, perhaps literally many millions. Although a high degree of precision could be built into the experiment, the information derived from it would be precise about the 10 "gamete equivalents," and it would be rash to generalize from such a narrow base. Any work on inbred lines refers, therefore, very strictly to those inbred lines only. Any conclusions from such work should not be deemed to apply to the species at large without supplementary evidence.

This conceptual restriction on the employment of inbred material is aggravated by its peculiar genetic composition. An inbred line cannot, by its very nature, contain any lethal genes; it is unlikely also to contain either semilethal or, as a result of sampling, rare recessive genes. Such genes will therefore not be found in a population of line crosses either. And yet lethal genes and rare recessives are an important feature of outbred populations, as they are of wild populations in the field. These considerations also detract from the usefulness of inbred material. To all this, we must add the disadvantage of peculiarities in gene frequencies. A cross between four inbred lines, for instance, means that no allele can have a frequency lower than 0.25 in the derived population. Yet, gene frequencies figure prominently in all our discussions of genetic parameters; changes in gene frequencies can radically alter all of them. The application of results from inbred lines and their derivatives to outbred populations should thus be exercised with extreme caution. An inbred line represents a unique and extraordinary situation in biology; line crosses are hardly less unique and extraordinary. An investigator, contemplating the employment of inbred lines, should therefore reflect seriously on the kind of information he hopes to obtain from them and ask whether it covers the range of variation in which he is basically interested.

The use of inbred strains has been mentioned twice in this chapter.

Firstly, it was said that differences between inbred strains kept under uniform conditions were evidence of genetic variation in a character. Secondly, and less certainly, it was suggested that isogenic material could be employed to assess the environmental component of variance. But beyond these two uses, the value of inbred strains in research work on quantitative characters must be questioned, unless interest rests on the strains themselves and on the peculiar combination of circumstances they represent.

SELECTION

We must now consider briefly the second agency whereby changes in gene frequencies may be brought about, namely, selection. One must distinguish, as in the case of inbreeding, between the usefulness of selection in the production of special strains and its usefulness as a research tool in quantitative genetics. In the former case, selection is a very valuable method; in the latter case, it has much less to contribute, for reasons that will be explained. Much of the theory of selection is devoted to the prediction of its results and to the evaluation of the efficacy of various methods of changing a character in a required direction. We shall consider here only a few of the salient features of the theory in order to illustrate the concepts involved.

Selection may be defined in terms of differential fertility, whereby individuals do not contribute equally to the next generation. Under experimental conditions, the population is usually truncated at some point on a phenotypic scale; the individuals on one side of the truncation point are allowed to breed, while those on the other side are not. The mean value of the selected individuals deviates from the population mean by a certain amount, S, termed the *selection differential*. This deviation is composed of genetic and nongenetic components, and the only part of it reflected in the mean performance of the progeny of selected animals is that due to the average effect of the genes. The relative importance of the average effect of the genes, compared with other sources of variation, was seen to be measured by the heritability. Therefore, the deviation of the progeny (R) from the population mean is given by

$$R = h^2 S$$

as derived previously. The deviation of the progeny is known as the *response to selection*.

This formula suggests another way of estimating the heritability. The selection differential and the response are both easily measured in practice, and from the relationship just given, the heritability can be calculated. To reduce sampling error, both S and R should be cumulated over several generations and plotted generation by generation. The slope of the least-squares regression line of R on S through these points then gives the heritability. The value so obtained should be termed the *realized heritability*, i.e., what is observed in practice.

One snag in this context is that, although the selection applied

truncates the population sharply, the selected parents will still differ among themselves in their contribution to the next generation. To correct for this, the selection differential of each parent should be weighted by the number of progeny it contributes for measurements. It is the figures so weighted that should be cumulated to arrive at the cumulated selection differential mentioned above. With some species, such as Drosophila, which provide many offspring, it may be more convenient to enforce equal representation among the progeny by taking random samples of equal size from each mating. Sterile matings must, of course, be excluded in calculating the selection differentials.

The selection differentials are not exactly equal for males and females, even when pair matings are employed; the two should be averaged for each generation. As the females tend to limit the rate of reproduction, the selection differential can often be increased by mating one male to several females and thereby achieve high differentials on the male side. The limit in this direction is usually the need to avoid excessive inbreeding by the restriction so imposed on the population size. In the case of laboratory mice, it has been found in practice that, provided the number of parents does not fall below the equivalent of 10 pair matings and if provision is made for each fertile mating to be represented in the next generation, the populations as a rule do not accumulate the more obvious effects of inbreeding. But if individuals are selected irrespective of the family from which they derive, the number of parents should be doubled.

The value of the selection differential depends on the proportion of animals selected and also on the phenotypic standard deviation of the trait. It is convenient to measure the selection applied in terms of the *intensity of selection* (i), which is in fact the selection differential measured in standard terms:

$$i = \frac{S}{\sigma_P}$$

where σ_P is the phenotypic standard deviation. By equating S to $i\sigma_P$, the formula derived previously becomes

$$R = i\sigma_P h^2$$

This shows how the response to selection depends on three factors: the intensity of selection, the phenotypic standard deviation of the character, and the heritability. The intensity of selection, expressed in this form, depends entirely on the proportion of animals selected, though the relationship is not a linear one. There are available tables, based on the normal distribution, that give the value of i for a given proportion selected.

Responses to selection, as is well known, tend to be rather erratic. Although the response in the desired direction may be quite apparent when the progress over several generations is surveyed, fluctuations from generation to generation render short-term assessments unreliable. These fluctuations may arise either through accidents of sampling or as a result of environmental changes. The first cause is beyond our control, and its influence can be estimated only by running

replicate selection lines concurrently. The second cause is all too frequently of unknown origin, but some of these environmental creases can be ironed out by the use of an unselected control population. The control and the selected lines should derive from the same base population and should be of identical structure. with the exception that in the control population the parents should be chosen and mated at random. The response to selection should then be measured as a deviation from the control line.

If separate and simultaneous selection is carried out for "high" and "low" expression of the character (two-way selection), one line acts as a partial control for the other. The effect of the selection can then be judged from the divergence between the two lines. This procedure, however, does not enable us to recognize a fairly frequent feature of selection programs, namely, the asymmetry of the response, which means that selection in one direction brings about a more rapid change in the mean than it does in the other. A behavioral example is discussed by Hirsch in Chapter 12. Without an unselected control line, the separate responses in the two directions cannot be adequately evaluated.

The pattern of the response to selection is itself of intrinsic interest. The subject is discussed in some detail by Falconer (1955). As mentioned, asymmetry of the response is common, and it may arise from a variety of causes, some of the more obvious ones being as follows:

1 Natural selection may oppose the artificial selection in one direction, while assisting it in the other.

2 Selection may favor heterozygotes in one direction, which of course segregate out homozygotes, thereby retarding the response.

3 The selection differentials attained in the two lines may not be equal, either as a result of the variances being different or because of differential fertility in the two lines.

4 If directional dominance affects the character, the line selected for the recessives will show a more rapid response.

5 Inbreeding depression may affect the mean of the character in the same direction in both lines, opposing one while reinforcing the other.

6 The characters may be of partially independent genetic origin. For instance, body weight may be increased by increasing the relative amount of adipose tissue, while selection in the opposite direction may soon reduce the fat to a minimum; progress in the low line would then become contingent upon a reduction in bodily dimensions.

Except in its pragmatic aspects, therefore, selection is not a potent method in genetic analysis. Short-term responses, if measured accurately, may act as a useful check on theory and on the accuracy of parameters estimated from some base population. In the absence of additive variance, the character will not respond to selection.

Continued selection invariably results, eventually, in the cessation of

the response, and a limit is reached beyond which no further progress is obtained. This *limit to selection*, also sometimes referred to as a "plateau" or a "ceiling" or other odd words, should be approached asymptotically as the additive genetic variance becomes exhausted. Occasionally, a renewed response may appear in practice, through the introduction of new variance in the form of a mutation of a major gene or the formation of rare recombinants. But usually a selected line at the limit remains at a fairly constant average level despite perhaps sharp fluctuations from generation to generation. If the additive variance has in fact been exhausted, then selection can do nothing to shift the mean level of the line in either direction. In practice, however, this is not always the case, and selection is often required purely in order to maintain the level of the line. When the selection is relaxed, the average performance often tends to revert toward that of the base population. This may indicate either the opposing force of natural selection or that the artificial selection had favored heterozygotes. Under these conditions, selected lines at the limit often respond rather rapidly to reversed selection, again indicating that the additive variance had not, in fact, been exhausted. When one adds the possibility of some genetic independence between the "high" and "low" expression of the character, one can appreciate that the possible complexities of the situation again defy precise genetic interpretation.

Because of these difficulties, the nature of the limits to selection has not yet been fully explored. No doubt the exhaustion of the additive genetic variance is often an adequate explanation. However, the response sometimes ceases too abruptly for this to be plausible; frequently a line responds more or less linearly and then suddenly stops. This obviously does not accord with a model of asymptotic depletion of the variance. As mentioned, natural selection or selection for heterozygotes often prevents progress while a considerable amount of additive variance remains. Indeed, the likelihood that natural selection is at work is often all too conspicuous in the form of widespread sterility or reduced fertility in selected lines.

Despite what may appear to some as a plethora of literature on selection work, there is still insufficient information to indicate the expected magnitude of the response to selection and how long the response may be expected to continue. A theory of limits published by Robertson (1960) suggests that, in terms of our formulation, the expected limit of selection is a function only of the product Ni, where N is the effective size of the population and i is the intensity of selection. Robertson shows further that one-half of the total response should be attained in $1.4N$ generations for genes that act additively and that this may approach $2N$ generations for rare recessive genes.

Selection is usually considered in terms of selecting individuals on the basis of their own phenotypic merit and is most conveniently discussed in such terms. This is usually called *individual* or *mass* selection. However, selection need not and does not always take this form. If, for instance, an individual must be killed in order to assay some hormone, it cannot be used for further breeding; it cannot itself be

selected, unless it has been possible to obtain and maintain offspring from all individuals liable to be selected, which would make excessive demands on facilities. Under these conditions, the stock must be propagated from the relatives, e.g., sibs, of "selected" individuals. In animal breeding, a phenotypic assessment of an individual is sometimes based on the performance of its progeny; e.g., bulls are selected on the milk yield of their daughters. Often information is available on the performance of relatives, and the question arises whether such information could and should be employed in assessing the genotypic merit of an individual. This is a big subject in its own right, though chiefly of relevance in the realms of animal breeding. In the laboratory, the question is usually simplified to the consideration of families of either full or half sibs. The question then is whether an individual should be selected entirely on its own merit, or whether it should be selected by its relative merit compared with other members of its family, or whether whole families should be selected on the basis of the family mean, without regard to individual deviations within the family. These three forms of selection are known as *individual* selection, *within-family* selection, and *between-family* selection, respectively. We shall not enter here into the algebra and statistics involved but merely indicate the main conclusions.

The general formula

$$R = i\sigma_p h^2$$

can be modified into

$$R_w = i\sigma_w h_w^2 \qquad \text{and} \qquad R_f = i\sigma_f h_f^2$$

where the subscripts w and f refer to within-family and between-family terms, respectively. Now, σ_w and σ_f can be expressed in terms of σ_p, and h_w^2 and h_f^2 in terms of h^2. By doing this, R_w and R_f can both be formulated in terms of $i\sigma_p h^2$ weighted by a term containing n, the number of individuals in the family, and t, the intraclass correlation. The relevant formulae are a bit complicated though not difficult to derive. When expressed in this way, R_w and R_f can then be compared with R, the response to individual selection, by substituting values for n and t. The intraclass correlation is the more important of the two, with the following general conclusion. Where t, and therefore the heritability, is low, the response from between-family selection exceeds the other two. Where t is high, and the heritability now may or may not be high, the maximum response is attained from within-family selection. But over a wide range of intermediate values of t, individual selection surpasses both of the more complicated methods. The critical values of t depend on n, the number in the family, and the reader is referred to Falconer (1960) for the exact formulae and their derivation.

A high intraclass correlation can result only from a proportionately large component of variance due to common environment, making full sibs similar. In these cases only should individuals be selected on the basis of their deviation from the family mean. The method has a disadvantage of utilizing only one-half of the additive genetic variance but

it has one redeeming feature. As each family is represented in the succeeding generation, the effective population size is doubled, as explained by Falconer (1960), and the method is therefore economical in space and facilities. Between-family selection, on the other hand, is costly in terms of facilities, as many more families must be measured than are selected, the remainder being discarded. It is therefore fortunate that individual selection is frequently the one that gives the maximum response.

In concluding this section on selection, we should note that it is possible to combine information on an individual's own merit with that on its family to arrive at an index of the expected breeding value. The individual's phenotype (P), expressed as a deviation from the population mean, can be regarded as a sum of two parts: firstly, as a deviation (P_f) of the family mean from the population mean and, secondly, as a deviation (P_{ic}) of the individual from its own family mean:

$$P = P_f + P_{ic}$$

The appropriate weighting factors for the expected breeding value are as follows:

$$Exp\ BV = h_f^2 P_f + h_{ic}^2 P_{ic}$$

This then becomes the index of selection. It might be added that *combined selection*, as this method is called, is seldom worth the trouble. Its superiority over some other form of selection is never very great and is often quite trivial.

CORRELATED RESPONSES TO SELECTION

Selection for any character may result in a concomitant change in some other character or characters. Such a phenomenon is referred to as a *correlated response* to selection, and we must examine briefly its genetic causation.

As a model, we shall postulate that selection for some character X changes also the mean level of performance of some other character Y. The connecting bridge is the *genetic correlation* (r_A), which may be defined as the correlation between the breeding values for the two traits, each measured in every individual. Thus, if the breeding values for X and Y are measured in a number of individuals, each individual provides a pair of observations which can be correlated. Such a correlation can arise only if the two characters are affected by "common genes"; in the long run, this implies pleiotropy, while in the short run a correlation may easily arise as a result of linkage; i.e., for "common genes," read "common chromosomal segments." Correlations of the latter kind can be particularly prevalent in crosses between divergent strains, though they fade through time as equilibrium is established between the coupling and repulsion phases of linkage.

The principle of the formulation is that the phenotypic covariance between two characters, measured on the same individuals, can be partitioned into an additive genetic component and a (nonadditive +

environmental) component, in a way strictly analogous to that previously employed for variances. If the breeding values themselves were correlated, as suggested in the definition, the additive genetic covariance alone would be derived. In practice, however, this usually proves too cumbersome, and the less direct method of partitioning is employed, by relating observational components of covariance to causal components. The reader is referred to Falconer (1960) for details. The genetic correlation between X and Y $[r_{(A)XY}]$ is then obtained by relating the additive covariance $[cov_{(A)XY}]$ to the two additive variances, according to the usual formula for a correlation:

$$r_{(A)XY} = \frac{cov_{(A)XY}}{\sigma_{(A)X}\,\sigma_{(A)Y}}$$

The correlated change in character Y, when the selection is for character X, will depend upon the regression of the breeding value for Y on the breeding value for X. This regression is the ratio of the additive covariance to the additive variance of X.

$$b_{(A)YX} = \frac{cov_{(A)YX}}{\sigma^2_{(A)X}} = r_{(A)}\frac{\sigma_{(A)Y}}{\sigma_{(A)X}}$$

The direct response in X was given as

$$R_X = ih_X{}^2\sigma_{(P)X}$$

where $\sigma_{(P)X}$ is the phenotypic standard deviation of X. This can be rewritten in terms of the additive standard deviation:

$$R_X = ih_X\sigma_{(A)X}$$

where h_X is the square root of the heritability. Hence the correlated response in Y, (CR_Y), can be formulated as follows:

$$\begin{aligned} CR_Y &= b_{(A)YX}R_X \\ &= r_A\frac{\sigma_{(A)Y}}{\sigma_{(A)X}}\,ih_X\sigma_{(A)X} \\ &= ih_Xr_A\sigma_{(A)Y} \end{aligned}$$

Since

$$h_Y{}^2 = \frac{\sigma_{(A)Y}{}^2}{\sigma_{(P)Y}{}^2}$$

$$\sigma_{(A)Y} = h_Y\sigma_{(P)Y}$$

so that

$$CR_Y = ih_Xh_Yr_A\sigma_{(P)Y}$$

The only purpose of this brief formulation is to show the number of factors that may influence the correlated response in one character when selection proceeds for another. The presence of a correlated response demonstrates only that none of the five factors in the formula is zero. The likeliest one to be zero is, of course, the genetic correlation. But it is much more difficult to argue that the absence of a correlated response means that r_A is, in fact, zero, especially if either of

the two heritabilities is low. And it can never be argued that, because a genetic correlation cannot be detected, the characters do not share some common genes. It is easily conceivable that one gene could increase one character and decrease the other, while another gene had the opposite effect, both genes thus masking each other. This underlines yet again the difficulty of drawing precise genetic conclusions from the responses to selection. The chief usefulness of correlated responses in genetic analysis is to act as a check on previous estimates of the heritabilities and the genetic correlation.

As a final remark in this section, it should be noted that, if the heritabilities are known, the formula given suggests a method of estimating the genetic correlation by measuring the correlated response. A better method, using selection data, is to select for each of the two characters separately and to observe the correlated response of the other character in each case. The formulae given can then be manipulated to show that

$$r_A^2 = \frac{CR_X}{R_X}\,\frac{CR_Y}{R_Y}$$

In accordance with our previous terminology, the estimate so obtained should perhaps be termed the "realized genetic correlation," to distinguish it from a priori estimates derived from the methods of partitioning the covariance.

CONCLUDING REMARKS

What I have attempted in this chapter is to present some basic concepts as they affect our thinking about the problem of quantitative variation in genetics and to indicate a few of the methods whereby the sources of such variation are explored. It has not been possible, in the space available, to develop some of these ideas very far; neither has it been possible to discuss at all the exploitation of the products of selection and inbreeding in genetic analysis. The latter objective would have been difficult in any case, as each situation demands a particular approach and a specific analysis, based on the investigator's understanding of the biology of the system with which he is working. Instead, I have concentrated on presenting the subject almost as an attitude of mind, at times, toward situations where genes are at work but where they cannot be identified individually. Very often, a lack of precision in the interpretation came to the surface. This is partly a reflection of the fact that I limited myself to certain objectives, but also it partly reflects the genetic system involved. For it is an accepted fact of life that genes hunt in packs. As such, the pack must be studied as a pack and this, by its nature, poses difficulties. While it is always of interest to learn about some of the quirks of isolated members of the pack, the action of these individuals, as individuals, is almost irrelevant for many purposes. They may have little bearing on the action of the pack as a whole. From this point of view, the quantitative aspects of genetic systems are fundamental and basic to many biological problems.

It will be obvious to most that the treatment in this chapter is far from being exhaustive; indeed, many important topics are not even mentioned. I should not like it to be thought either that the approach adopted is an exclusive one; I have employed the terminology, symbolism, and formulation with which I am most conversant. I should like to believe, however, that the concepts developed here are basic to any approach.

Much of the stimulation of quantitative genetics derives from theoretical studies, developed largely from the works of R. A. Fisher, J. B. S. Haldane, and Sewall Wright in the 1920s and 1930s. Some of these early theoretical papers can still be claimed to be the cornerstones of the subject. Primarily, however, quantitative genetics should be regarded as an empirical science, and it is from more experimental work, especially perhaps with characters as yet uninvestigated, that further impetus and progress should be expected.

PAPER 26.

Some evolutionary implications of behaviour.

Canad. J. Genet. Cytol. 11, 419-435. 1967

by

R.C. ROBERTS

Reprinted in Canada from CANADIAN JOURNAL OF GENETICS AND CYTOLOGY, Journal Canadien de Génétique et de Cytologie, Volume IX, Number 3, September 1967

SOME EVOLUTIONARY IMPLICATIONS OF BEHAVIOR[1]

R. C. ROBERTS[2]

Department of Psychology and Institute for Behavioral Genetics, University of Colorado, Boulder, Colorado

Introduction

"Many instincts are so wonderful that their development will probably appear to the reader a difficulty sufficient to overthrow my whole theory. I may here premise that I have nothing to do with the origin of the mental powers, any more than I have with that of life itself."

It was thus that Charles Darwin meekly opened his chapter on "Instinct" in *The Origin of Species.* Though Darwin's modesty could well be commended to some latter-day students of evolution, many people today would regard his diffidence and veneration of "the mental powers" as perhaps a little unnecessary. Be that as it may, it is certain that behavior is an aspect of evolution that cannot be neglected, for as Darwin himself wrote a little further in the same chapter:

"It will be universally admitted that instincts are as important as corporal structures for the welfare of each species, under its present conditions of life. Under changed conditions of life, it is at least possible that slight modifications of instinct might be profitable to a species; and if it can be shown that instincts do vary ever so little, then I can see no difficulty in natural selection preserving and continually accumulating variations of instinct to any extent that was profitable."

We can probably substitute the word "behavior" for "instinct" in the above paragraph without drastically affecting Darwin's original meaning. In any case, the distinction between instinct and the inherited capacity to learn a behavior is, at best, blurred and arbitrary. It seems clear, therefore, that the implications of behavior in evolution have been recognized since the theory of evolution first assumed its prominence in biological thinking. The animal's quest for sustenance, its immunity from predators and other hazards, its ability to reproduce and successfully rear its young all patently involve various aspects of its behavior. An animal's behavior must therefore be an important component of its natural fitness.

Just how important behavior may have been is illustrated by some anthropological evidence on human evolution. A useful review of some of the salient features is provided by Washburn (1960). It was apparently thought until quite recently that man had evolved to an advanced state before he developed any capacity to use tools. By now, however, it seems certain that some much more primitive forms, the man-apes, had learned to make and use crude stone tools earlier than 500,000 years ago, which precedes any

---

[1]Substance of Annual Invitation Lecture given to the Genetics Society of Canada, at Montreal, June 13, 1967.

[2]Permanent Address: A.R.C. Unit of Animal Genetics, Institute of Animal Genetics, Edinburgh 9, Scotland.

"Preparation of this paper was supported in part by National Institutes of Health grant number GM 14547."

Manuscript received July 31, 1967.

Can. J. Genet. Cytol. 9: 419-435. 1967.

skeletal or cultural evidence for the existence of the genus *Homo*. These man-apes had brains no larger than those of apes now living, and they were not fully bipedal. Washburn argues that the use of tools among these creatures had profound influences on their further evolution. The use of tools favoured the development of fully bipedal locomotion, to release the forelimbs for other functions. The shape of the hand was modified to manipulate the tools more efficiently, especially with respect to the use of the thumb. This in turn led to an enlargement of cortical areas in the brain to cope with increased sensory and motor functions associated with a redesigned hand. The enlargement of the head size had obstetrical consequences which led to the human child being born at an earlier stage of development than in other primates, but this became possible only because a fully bipedal mother had arms and hands to hold an immature, non-clinging infant. There are many other fascinating aspects of this story, but perhaps enough has been suggested to indicate that one piece of behavior – the initial use of tools – not only featured in human evolution but indeed rendered that particular evolution possible. It should also be noted that the evolution of behavior and of "corporal structures" continually interact, a further modification of one being contingent upon some prior modification of the other, as is well illustrated in Washburn's review. Some current writings on "cultural evolution" – as if it had a non-organic basis – seem to ignore the fact that natural selection moulded the biological entities on which social and cultural attributes are based.

It is not my purpose here to examine the evolution of behavior from a comprehensive viewpoint. The reader is referred to Manning's (1963) excellent review of the broader aspects of the topic. I want to limit myself to a much narrower objective. The study of behavior in the laboratory usually involves measuring some aspect of behavior in a particular apparatus. Thus, when investigators talk of "activity" or "learning", they must always define their characters operationally in terms of some measurement. While it is easy to imagine that maternal care of the young, say, has been subject to strong natural selection, it is difficult to imagine a laboratory measurement that would adequately define "maternal care" in a broad sense. The behavioral measurements that experimenters are forced to use may therefore reflect only obliquely and dimly the behaviors that featured significantly in the evolution of the organism. A measurement of activity in an arena, say, may be an inadequate representation of the activity of the animal's distant ancestors in the wild.

It is proposed in this paper that information on the genetical architecture of a behavioral measurement may be employed to identify what aspects of behavior may have been important from an evolutionary perspective, and what aspects may have been remote from the forces of natural selection. The genetical basis of the proposed method is examined in the next section; readers versed in quantitative genetics should be warned that there is nothing particularly new or original in this part. That section is followed by a preliminary attempt to apply the method in practice.

The Evolutionary Moulding of a Metric Character

The central concept in any formulation of quantitative genetics is the *additive genetic variance* – the variance in the population due to the average (or additive) effects of segregating genes. The additive variance in a character is most conveniently expressed as the *heritability* of the character, which is the

proportion of the total phenotypic variance attributable to additive genetic sources. Because of its importance for both descriptive and predictive purposes, the heritability has been determined, often repeatedly, for a range of characters in both laboratory and farm animals, and the results are sufficiently consistent for the following general conclusion to emerge. Characters that are closely connected with the natural fitness of the organism tend to have low heritabilities – characters like conception rate, litter size or measures of juvenile viability. Characters that are less obviously directly related to fitness – like measures of size, for instance – seem in general to have intermediate heritabilities, while those of trivial evolutionary significance – white spotting in modern breeds of dairy cattle seems to be the favorite example – have very high heritabilities.

The genetic properties of a population as we find it today must be ascribed to the cumulative evolutionary history of the organism. Thus, natural selection constantly acts to improve fitness, and in a population at equilibrium, there can be no additive variance left in fitness itself. It will have been exhausted by natural selection. There is no single laboratory or field measurement to which the term "fitness" can be applied, but it appears logical that the amount of additive variance displayed by any character must reduce, the closer it becomes to fitness itself. It is in this way that the general conclusion just given is usually explained.

This presumed relationship between the genetic properties of a metric trait and natural fitness, as propounded above, is a rationalization of empirical observations. The theoretical basis of the concept has been the focus of less attention. Let us regard any character (X) as being composed of two parts: firstly, a part ϕ, any variance in which is defined as part of the variance in fitness, and sceondly, a remainder (r), which is defined as being uncorrelated with ϕ. Thus

$$X = \phi + r$$

and since ϕ and r are uncorrelated

$$V_X = V_\phi + V_r$$

where V is the variance whose source is denoted by the subscript. This partitioning also extends to the additive genetic variances, $V_{(A)}$;

$$V_{(A)X} = V_{(A)\phi} + V_{(A)r}$$

To illustrate that such a partitioning may be a realistic concept, let us reflect momentarily on the connection between, say, body size and fitness. It appears likely that, over a certain range of observations, larger members of a species may be more fit than the smaller ones, and for numerous reasons. For instance, larger females tend to shed more eggs and rear bigger litters, while the larger males probably compete more successfully for territory, food and mates. Any variance in body size associated with such advantages (or disadvantages) could therefore be termed the part ϕ in the above partitioning. However, it does not follow that all causes of variation in body size contribute equally to fitness. It is easy to imagine that fat deposits, for instance, within limits affect neither ovulation rate nor competitive ability. Much variation from such causes is therefore essentially neutral with respect to fitness, and represents the part r in the partitioning.

The component $V_{(A)r}$ could equally well be regarded as the residual additive variance in a character if there was a range of values in X over which, for one reason or another, all individuals were equally fit, but that deviants from this range, in one or both directions, had been selected against. Any additive variance due to these deviants could then be termed $V_{(A)\phi}$. The deviants could be exterminated in one of two ways. Either loci directly responsible for them could be fixed for less extreme alleles, or a sufficient number of loci might be fixed at random to permit only a restricted range due to residual segregation. If the latter were the case, crosses between populations randomly fixed for different alleles might be somewhat less fit, on average, than the parent populations.

Under natural selection, as the additive genetic variance in fitness is depleted, $V_{(A)\phi}$ must obviously disappear with it. Any additive genetic variance left in the character — what was conceptually designated $V_{(A)r}$ — is thus an inverse reflection of the connection between the character and fitness. A large additive genetic component implies a remote connection with fitness, while characters close to fitness will have had most of their additive variance depleted by natural selection. The argument can be emphasized by considering an extreme case, as follows. Suppose that *all* of the variance in X is part of the variance in fitness, as it would be for instance in an artificial selection experiment. The connection between the character and fitness may then be formulated by regarding fitness (F) as being composed of two parts: firstly, a part associated with character X, and secondly, an uncorrelated remainder (Y), which represents all other attributes of the organism that affect fitness.

$$F = X + Y$$

The additive genetic covariance between X and F is then:

$$cov_{(A)XF} = cov_{(A)}[X(X + Y)] = V_{(A)X}$$

since cov_{XY} was defined as zero. As the additive variance in fitness is depleted then so must the additive covariance with X, and thus the additive variance in X, also become depleted, until all vanish simultaneously. It was shown by Falconer (1966) that what is termed here $cov_{(A)XF}$ represents the response in character X under natural selection. By expressing the covariance, as we can do in this special case, as the additive genetic variance in X, the treatment falls into a familiar formulation in quantitative genetics. The formulation is also strictly analogous to Fisher's (1930) "Fundamental Theorem of Natural Selection," which states that the change in fitness is determined by the additive genetic variance in fitness. What we have been discussing all along is, in fact, a necessary consequence of Fisher's theorem.

While at least some part of the additive genetic variance is likely to become exhausted under natural selection, it is clear that the non-additive components, arising from dominance and interaction, are not subjected to direct depletion. It should be noted, however, that these components will be affected by changes in gene frequencies resulting from the selection, and that if the gene frequencies are driven to fixation, then no genetic variance of any description will be left in the character. But it was suggested earlier that complete fixation through natural selection will not occur unless all of the variance in a character is related to fitness. To the extent that this contingency is improbable, except for fitness itself, complete fixation is therefore unlikely.

It is therefore not surprising that natural populations, whenever they are subjected to experimental analysis, reveal at least some evolutionary reserves in the form of latent genetic variation. We have already discussed the additive genetic component of this variance, and we should now see whether there is anything we can deduce about the relationship between non-additive genetic variance and natural fitness. Two aspects of the topic will have to be considered simultaneously. Firstly, the amount of non-additive variance displayed by a character will be determined very largely by the amount of dominance shown by the genes affecting the character; in order to simplify the discussion, I shall trust that the dominance variance is an adequate indicator of all non-additive variance. Secondly, the non-additive variance is a function of gene frequencies, as mentioned above, and the effect of natural selection on this factor must also be considered.

If, as Lerner (1954) suggested was possible, natural selection favours heterozygous individuals, then this is effective overdominance with respect to fitness. A character closely connected with fitness might be expected therefore to reflect the same phenomenon of overdominance. Should this be the case, then the gene frequencies at the relevant loci will equilibriate at intermediate values, the exact value depending on the relative disadvantage of the two homozygotes compared to the heterozygote. But the important point for this discussion is that at equilibrium gene frequencies, all of the genetic variance will be non-additive. Furthermore, the amount of dominance variance will be magnified by the intermediacy of the gene frequencies. Even if only a few of the loci affecting the character exhibited overdominance, the non-additive genetic variance due to such loci could constitute a swamping proportion of the total genetic variance.

It is only fair to add that overdominance, even if its likely occurrence is admitted, is not an easy phenomenon to establish in practice. Fortunately, perhaps, for our limited and immediate purposes, this is unimportant, for alternative genetic postulates lead to similar expectations with respect to the partitioning of the genetic variance into its various components. Robertson (1955) reasoned that the effects of genes on characters closely connected with fitness would show more dominance than the effects of genes further removed from fitness. The genes most favoured by natural selection should tend therefore to be those with a dominant expression, or perhaps those whose expression would evolve, through time, to be dominant. Furthermore, this dominance, with respect to a particular character, should be preponderantly in the one direction. The effect of this mechanism on various genetic components is unambiguous. As the frequencies of the dominant genes – those close to fitness – are increased, the dominance variance, through decreasing in absolute terms, forms an ever-increasing fraction of the total genetic variance. The reader is referred to Falconer (1960) for a fuller discussion of the relationships between gene frequency and the proportions of various components of genetic variance.

There seem therefore to be alternative theoretical approaches which suggest that the closer a character is to fitness itself, the more non-additive variance it should display. But irrespective of theoretical considerations, there is a plethora of experimental evidence which accords with this expectation. The characters most susceptible to the effects of inbreeding – or by the same token, showing the greatest amount of heterosis – are those which by other

criteria are the most obviously and closely connected with the organism's natural fitness. Such characters are various measures of fertility, like conception rate, litter size or prepubertal viability. Since non-additive genetic variance is the prerequisite of inbreeding depression, the implication is clear that those characters which seem to be closest to fitness also seem to feature non-additive genetic variance most prominently.

We have therefore a congruity of theoretical considerations and experimental findings which lead to a clear conclusion. Characters close to fitness display little additive but much non-additive genetic variance. Conversely, characters remote from fitness have the proportions reversed. There is, however, an important exception which bears on the utility of the conclusion as it will be applied in a moment. The exception relates to characters that have intermediate optima with respect to fitness. The reader is referred to Robertson (1956) for a more comprehensive treatment of this subject. Robertson recognizes that an intermediate optimum may arise from various causes. In some cases, this will not prevent fixation in the usual manner, but if the intermediate value is imposed by a correlated character, then the additive genetic variance may not become exhausted by natural selection even though the character is related to fitness. This is perhaps best illustrated by a specific example. Manning (1961) selected for mating speed in *Drosophila melanogaster,* and though the character selected would apparently be an important aspect of the organism's fitness, Manning found a ready response which, over the early generations, corresponded to a heritability of 0.30. Of particular relevance, however, was the correlated response to selection observed in his high lines; the locomotor activity of the fast mating lines was reduced. Thus we may suppose that under natural selection, mating speed had been maintained at a submaximal level, since a further increase would not compensate the fly for the consequent reduction in activity. The fittest fly is therefore the one with the best "balance" between mating speed and activity, and though both characters affect fitness in a fairly direct manner, additive variance will remain in both. Ewing (1963) in fact confirmed the presence of additive variance in activity, at least on some measures. We can imagine that over a certain range of intermediate values, some reduction in activity could be compensated for by an increase in mating speed, and *vice versa;* but equally, that there are limits to this tolerance, and no amount of locomotor activity would compensate for the failure to mate.

Despite complications of this nature, the general relationship between the genetic architecture of a character and its role as a determinant of fitness is sufficiently firm and uncontentious to suggest its possible utility in an unexplored situation. The suggestion is that by examining the genetic profiles of various characters, we ought to be able to gain some insight into the evolutionary significance of the character. This procedure reverses the usual sequence of thought, whereby we consider the implications of the character in Darwinian terms and then rationalize its genetics properties; this can be done with many morphological traits without any excessive mental contortions. But with behavioral measurements – certainly the laboratory measurements – the task is much more difficult. There seems to be enough disparity between the behavior observed from one piece of apparatus to another – superficially still measuring activity, learning or whatever – to make one suspect that the laboratory measurement is specifically related to the apparatus employed, and

the reaction of the animal to it. It therefore becomes a tenuous exercise to try to relate that behavior to natural fitness since this involves a subjective appraisal of the relationship between the apparatus and behavior in the "wild." For this reason, I suggest that the examination of the genetic properties of the character may relate that character to natural fitness in more reliable and objective terms. I want to stress that I am in no way reflecting on the operational validity of the laboratory measurements. But the evolutionary moulding of a character does have far-reaching implications for its genetic or biological interpretation. To the extent that this kind of interpretation is now the declared goal of many behaviorists, any attempt to pinpoint more accurately the evolutionary role of behavior may be worth while.

In the remainder of this paper, I want to explore the potential utility of this approach, merely to see what kind of conclusions may emerge. No claim is made that the treatment is in any sense a comprehensive one. The data have been chosen merely to illustrate the methodology. The establishment of unequivocal conclusions should probably await more and better data, and particularly, the discussion of such data by investigators better equipped than I am to examine them critically from a behavioral point of view.

An attempt to identify the evolutionary significance of some behavioral phenotypes

One of the most potent methodologies for investigating the genetic architecture is to inbreed some animals over a few generations, and record changes in the character under observation, or in as many characters as may conveniently be measured. A single inbred line makes minimal demands on space, so that in an organism like the mouse, the procedure is readily replicable on a large scale without straining the resources. Furthermore, lines soon begin to die out through infertility, and after 5 to 10 generations, the program usually brings itself to a natural conclusion. Unfortunately, however, no behavioral study seems to have been organized to determine the effects of inbreeding on the behavioral phenotypes. Many studies, on the other hand, have used existing inbred strains of mice for the standard crossing schemes, but for present purposes, such data are of little value. Strains of mice with long history of inbreeding are presumably the rare survivors among the very large number of strains that must have been started. The vast majority of these would have succumbed to the usual deleterious effects of inbreeding, and the survivors represent unique and peculiar circumstances. It is not surprising therefore that crosses between long-established inbred strains reveal a variety of undirected outcomes reflecting the idiosyncrasies of the parental strains. This is true even of a character like litter size (Roberts, 1965), although the effects on litter size of inbreeding an outbred population are notoriously, but universally, deleterious. Crosses between long-established inbred strains all too frequently fail to reveal the directional dominance which characterizes litter size in the mouse, as indeed it does in a vast array of other organisms. As the character under consideration becomes further removed from natural fitness, then qualms about the generality of results from inbred crosses cause increasing discomfort.

Because of these considerations, the drawing of illustratory material, for present purposes, is precluded from a disquietingly large proportion of behavioral genetics studies, because they have been geared to the crossing of inbred strains of mice. As far as animal work is concerned, we are left,

mostly, with some selection studies. While these frequently reveal the presence of additive genetic variance, it is usually not possible to derive any estimates of heritability from the published data. Where estimates of heritability are available, they tend to be imprecise for the usual statistical reasons. I shall therefore frequently employ subjective judgements to illustrate the general approach, knowing that errors in judgement may be revealed as more precise data becomes available.

(1) *Selection on measurements of learning*

Probably the most widely known and often quoted selection experiment on animal learning is Tryon's (1929, 1940, 1942) mammoth study. Following Tolman's (1924) pioneering effort, Tryon selected rats for maze-learning ability. After 8 generations, Tryon had produced two populations with virtually no overlap between them, with respect to the errors performed in running the maze. By any criterion, this must be judged to be a rapid response, and the conclusion must be that Tryon had started with a population in which there was considerable additive genetic variance in the character selected. It is possible, therefore, that Tryon had worked with a measurement that gave a poor reflection of any kind of learning ability which had been subjected to natural selection. However, much supplementary work has been done on Tryon's maze-bright and maze-dull rats which, cumulatively, suggests an alternative conclusion. All of these studies point to differences between the "brights" and the "dulls" in the details of their responses to selection. For instance, Krechevsky (1932, 1933) found that while the bright animals made much use of spatial cues, the dulls preferred visual cues. Wherry (1941) found that three components of learning – forward-going tendency, food pointing and goal gradient – could be employed to distinguish the pattern of learning in Tryon's dull and bright rats. Searle (1949) showed that the brights were not uniformly better on all learning tasks but that Tryon had probably selected for and against a "brightness" specific to his maze, or at least to a certain class of mazes. Later studies which might suggest a fairly general superiority to Tryon's original bright rats in a variety of other mazes are reported by Rosenzweig *et al.* (1960), Fehmi and McGaugh (1961) and McGaugh *et al.* (1962). The interpretation of these studies, however, is somewhat confounded by the report of Rowland and Woods (1961) who, after rebuilding Tryon's maze, found that what was originally the dull strain performed significantly better than the bright one. It is, however, only fair to add that this last study was based on very small samples.

The point of all this, for our purposes here, is that Tryon's rats when selected for maze-learning ability showed many concomitant but differential changes in subunits of the character and indeed other aspects of behavior (see, especially, Searle, 1949). This suggests that Tryon had selected for a complex of subunits that were intercorrelated with one another. If we were to suppose that there were intermediate optima for some of these subunits, we have the mechanism for the preservation of much additive variance in the "character" even though it had been subjected to natural selection. This interpretation has a subjective appeal over the notion that learning ability, even in a laboratory maze, is unconnected with the rat's natural fitness. But without a fuller account of the genetic properties of the character, speculatory interpretations are of limited value.

Tryon's study has been singled out for special comment partly because

of its historical prominance in behavioral genetics. Other selection experiments for learning in the rat could equally well have been discussed. Heron (1935, 1941) selected for maze-learning ability in a maze that was broadly similar, in principal, to Tryon's. He obtained an immediate and pronounced response. Thompson (1954) reported a two-way selection experiment in a Hebb-Williams maze. Briefly, this maze is one that can be made increasingly complex for successive trials. Again, good responses were observed in both directions. Bignami (1965) describes a selection experiment whereby rats had to avoid getting an electrical shock by running to the opposite side of a shuttle-box within 5 seconds of the onset of a light stimulus, or "warning." The selection produced a clear distinction between good and poor "avoiders" over 5 generations. All of these reports have been accompanied by supplementary studies on the nature of the response, and without repeating any of the argument, exactly the same general points could be made as were given in the discussion of Tryon's results.

It thus appears that much additive genetic variance remains in these laboratory measurements of the rat's learning ability. It is suggested that subunits of the measurements employed have probably been characterized by intermediate optima on an evolutionary time scale, and that furthermore, these optima have been imposed by correlated characters. Suppose, following Krechevsky's (1932, 1933) findings, we postulate that good maze-learning ability demands particular attention to spatial cues, and that this, for some reason, is negatively correlated genetically with attention to visual cues, which is a characteristic of the poor maze-learners. It is almost self-evident that a high natural fitness requires an adequate performance on both criteria, and that therefore, the genes controlling this complex will not be driven to fixation. A few such mechanisms would retain ample reserves of additive genetic variance in almost any laboratory measurement of learning; almost any measurement, but perhaps not all. There are reports of two selection studies, one by McDougall (1938) and the other by Kuppusawny (1947), where the results were entirely negative. In the behavioral genetics review literature, both of these studies are usually given short shrift and the negative results are received rather disparagingly. McDougall is supposed to have been grinding a Lamarkian axe, while Kuppusawny just failed to explain his techniques to everyone's entire satisfaction. However, to be fair, if we are going to be pernickety on procedural grounds, then most of the early selection works, and not just the behavioral ones either, fail to meet present-day standards of rigour.

Taking the negative results at face value, merely to illustrate a point, McDougall selected for 24 generations for poor performance on a brightness discrimination task in a water tank; he also selected for 10 generations for good performance. Kuppusawny selected for 10 generations in a water maze also. No response was observed in either experiment. While the genetical techniques were admittedly sloppy, other investigators also with sloppy tchniques produced a divergence. But it does not take much reflection on the natural habitat of the rat to suggest that brightness discrimination while swimming in water may indeed be an important factor in survival. If we can accept this, then it may well be that McDougall and Kuppusawny chose measurements in which there was little or no additive genetic variance left. There is support for this view in a paper by Agar *et al.* (1948). In the course of an extensive repeat of McDougall's experiment, conducted to disprove his Lamarkian

hypothesis, Agar *et al.*, calculated that the parent-offspring correlation on the learning criterion was only $+ 0.06 \pm 0.09$. They also did some retrospective "paper selection"; had this been applied during the course of the experiment, the evidence was that no response whatever would have been observed. Should this kind of finding be established in further studies, then we shall begin to know a little more about the evolution of learning, particularly why some aspects of learning may be more heritable (in the narrow sense) than others.

(2) *Selection on measurements of activity*

Activity is perhaps second only to learning as a domain of interest in experimental psychology. The first selection study on activity was reported by Dawson (1932). Dawson must in fact have been quite an innovator; not only did he think of selecting, but he departed drastically from the psychological practice of the time by using the mouse instead of the rat. However, he did not exploit this particular innovation very well, as he seems to have stopped selecting after 4 generations. Nevertheless, by selecting for "wildness" and "tameness", as measured by running time down a straight alley, he did succeed in making animals tamer, though he did not, over his short experiment, increase wildness.

About the same time, Rundquist (1933) published the results of the first 12 generations of selection for wheel-running activity in the rat. Though scaling difficulties make direct comparisons impossible, his responses were less impressive than those in the early learning selection studies. The main features of his results were that activity increased hardly at all, but that there was a marked response in the inactive line. Brody (1950) produced a second inactive strain to replace Rundquist's original low line, which had died out through infertility after the 25th generation. Brody claims that the second selection experiment produced results similar to the first.

Rundquist, in his paper, had already noted a decline in the fertility of his inactive line, even by generation 12. Not only did the percentage of fertile matings decline but also the interval between mating and parturition increased, and litter size in the inactive line was consistently smaller. If natural selection operated to this extent even under the favourable laboratory conditions of pair mating, it seems reasonable to assume that it would have applied *a forteriori* during the evolutionary history of the rat. This then relates activity to natural fitness in a fairly direct manner, and the failure of both Dawson and Rundquist to increase activity by selection occasions no surprise. There was, however, still some additive genetic variance left in the character, otherwise it would not have been possible to decrease activity. Genetically, this must be interpreted to mean that the optimum level of activity is imposed by a correlated character, as discussed in the preceeding section. This, however, does not explain the asymmetry of the response, though several possibilities are open. For instance, Rundquist's mating system, quite apart from the small sizes of his populations, permitted a moderate amount of inbreeding. Should inbreeding depression make both lines less active, and this is at least plausible, then it may be an adequate explanation. This would require some directional dominance in activity, which again links the character to natural fitness.

The only other selection experiments for activity have been done with *Drosophila*. Since activity in the fruit-fly and in the rat may mean quite different things, we need not necessarily expect them to be related biologically or evolutionarily. Ewing (1963) measured activity by selecting the first flies

to traverse a series of six glass tubes, where back-tracking was not permitted, and selecting also, as the inactive line, those that remained in the start box. A conspicuous response was found over eight generations, but further tests revealed that the response was, at least in part, apparatus specific. Particularly, Ewing's high and low lines showed no difference when tested singly in an arena. This suggested that the flies had been selected primarily for their "reactivity" to each other. Ewing tested this by repeating his experiment, only this time he ran flies singly in his original apparatus. There was an immediate but short-lived response for inactivity, whereas selection for higher activity yielded little if any change over 10 generations. This situation is strongly reminiscent of Rundquist's rat study, quoted earlier.

The implications of Ewing's study, for our purposes here, are far-reaching. Ewing has shown that the manipulation of some of the experimental variables can profoundly affect the genetic outcome. This suggests a technique that, if used intelligently, can contribute greatly to our understanding of what aspects of behavior may have been important on an evolutionary time scale. Though the behavioral implications are not clear to me, it seems that the high heritability found in Ewing's first experiment, and the much lower one in the second, indicates that group activity may have been a less important feature of natural fitness than individual activity. This is exactly the kind of information we need for the further exploration of the evolution of behavior.

A recent paper by Connolly (1966) describes another selection study for activity in *Drosophila,* this time in an arena. Connolly's experiment attains a degree of genetical sophistication uncommon in behavioral studies. In accordance with expectation based on a heritability of 0.51 ± 0.10, derived from the regression of offspring on parents in an unselected stock, Connolly produced high and low lines, each of which diverged from the control population. The heritability must be regarded as high, suggesting that spontaneous activity within the confines of an arena may not correspond very closely to any aspect of natural fitness. This suggestion should arouse little astonishment, since flies in the wild must move in three dimensions and in response to a whole galaxy of stimuli.

Though they are not entirely appropriate under the rubric of activity, it is convenient to mention here the several selection experiments for taxes in *Drosophila.* These experiments owe much to the elegant mass-screening techniques used by Hirsch and his collaborators. Hirsch and Boudreau (1958) reported a highly successful two-way selection for phototaxis, with an estimated heritability of 0.56. Hadler (1964a,b) reports similar results with a somewhat modified apparatus; the heritability in this case was also about 0.50. The spectacular responses for geotaxis, described by Erlenmeyer-Kimling *et al.* (1962) and by Hostetter and Hirsch (1967), are well known and point to an enormous amount of additive genetic variance in the character. From this, it might be suggested that the taxes criteria employed in the laboratory are far removed from any indicator of natural fitness. Yet, some supplementary experimentation requires that such a conclusion should be qualified. Even after a long period of initial selection, reversed selections for the opposite response in these studies have been astonishingly successful, pointing to the retention of a large amount of the original additive variance. This is well illustrated by Dobzhansky and Spassky (1962), whose high and low geotactic lines, after diverging for 18 generations of selection, converged to base level

after only 6 generations of reversed selection. Though some experiments on the relaxation of selection in the selected lines, as distinct from reversed selection, would have added useful information, we are driven to the firm conclusion that natural selection opposes the artificial selection in either direction. While the amount of additive genetic variance for such a situation is unexpectedly large, what we clearly have in the case of both photo- and geotaxis are characters with intermediate optima. It would, after all, be a singularly nonadaptive kind of behavior if a fly in nature persisted in the direction of its response to either the sun or gravity. The lability of these responses, in relation to other demands made on the fly, are presumably a necessary condition of survival. It now seems that this lability has a large additive genetic component, though this result would not necessarily have been expected.

What has emerged from this section is that different measures of activity are variously related to natural fitness, as deduced solely from the genetic properties of the measurement. To the extent that this accords with general expectation and can be related to behavioral considerations, the proposed approach inspires increasing confidence.

(3) *Selection on measurements of reactivity*

There are two major selection studies in this area which will not be discussed in great detail, since the main points of relevance to the present treatment have already emerged in a different context earlier. Hall (1938, 1951) selected rats bidirectionally for defecation in a brightly-lit open field, as a measure of emotionality. While he observed little response on low-defecation scores, he obtained a modest response in the other direction, which ceased after about 9 generations. The difference between his high and low lines amounted to less than three times the original phenotypic standard deviation, though inbreeding during the course of selection must have reduced Hall's potential divergence.

The second important selection study in this area was summarized by Broadhurst (1960). He selected rats, again in two directions, in an apparatus similar to Hall's, though on a somewhat different defecation measurement. Broadhurst prefers the less-loaded term "reactivity" to "emotionality." Again, a divergence was produced between the lines, this time the response being more pronounced in the downward direction. While Broadhurst's "reactive" line showed a slow and unsteady increase over the 10 generations he first reported, the "nonreactive" was approaching the asymptote of no defecation (during the testing period, of course) in 5 generations. In a detailed genetic analysis, Broadhurst computed estimates of various genetic parameters which, unfortunately for our purposes but in no way reflecting on Broadhurst's approach, cannot be translated readily into the terms of our earlier discussion of genetical considerations. But reverting to the subjective approach applied to Hall's data, the divergence between Broadhurst's lines amounted to about twice his original phenotypic standard deviation.

The two studies taken together demonstrate very clearly the presence of some additive genetic variance in measures of reactivity. The patterns of response and the total divergence suggest that measures of heritability would yield moderate rather than high estimates. This may indicate yet another case where natural selection has favoured an intermediate optimum, this time for reactivity, the intermediate level being imposed by a genetically correlated character. Broadhurst's work, in fact, provides a beautiful example of the way

in which this may have been brought about. Though inbreeding was a confounding genetic variable in his study, Broadhurst did not find any major decline in fertility, but what he did find was a drastic reduction in maternal care, in *both* lines, as measured by the percentage of deaths between birth and weaning. Should it prove to be true that deviations in reactivity in either direction adversely affect maternal performance, then this mechanism alone may be sufficient to retain a substantial amount of additive genetic variance in the character.

(4) *The genetic architecture of components of human intelligence*

Much effort has been devoted over the years to the measurement of human intelligence, with considerable attention being given to identification of factors that affect the measurement. "Intelligence" is, of course, a gross term which usually and most conveniently, is defined operationally as a score on some test. Performance within some educational system may be employed as a validation of the tests. The construction of these tests, which is an extensive topic in its own right, seems to have led (perhaps fortuitously) to measurements which often have strong genetic components. For instance, a review by Huntly (1966) indicates that somewhere between 60 and 90 per cent of the variance in intelligence, as estimated by predominantly verbal tests, can be attributed to genetic causes. Let us assume on general grounds, though the genetic analyses summarized by Huntly do not in particular warrant the assumption, that much if not most of this genetic variance is additive. The question then arises whether our culture now places a premium on attributes that may have been unimportant over most of man's evolutionary history. The complete answer to this question, should that become feasible, would have implications beyond its obvious heuristic value, since it may indicate the scope for the further evolution of human mental abilities. Though it may be premature as yet to consider, as for instance Huxley (1963) does, the deliberate direction of that evolution by eugenical means, it is still as well that such proposals should be argued on firmer bases of fact. McClearn (in press) emphasizes the need for information in this area. It is of interest therefore whether any aspects of intelligence have had their genetic variance exhausted by natural selection, and as a corollary, what aspects have retained the potential for further genetic change. It should be noted, however, that on the best evidence available, little genetic change will occur while the present fertility patterns (in relation to IQ scores) remain as they are. Falconer (1966), discussing this question, concluded that if there is any genetic change in intelligence, then that change is towards an increase in IQ score at the rate of one- or two-tenths of an IQ unit per generation. From this, it follows that the genetic architecture of IQ can not have been affected by the recent social consequences that revolve around educational performance.

What is of interest then, in the present context, is any information on the genetic structure of different components (or measurements) of "intelligence." One subdivision of intelligence that may be useful from this angle is Thurstone's (1941) Primary Mental Ability (PMA) test. This test, which derives from a factor analysis approach, gives six measures (PMA's) to which various names (listed in the table below) are given. It does not matter for our purposes whether the term "reasoning", for instance, corresponds to the common usage of that term or not; the data stemming from the test are unaltered even if the labels were to be changed. If some of them have little additive variance

while others have more, then we obviously have a subdivision of intelligence that goes some way, at least, to meet our present requirements.

Three studies have been reported which indicate the extent to which various PMAs are inherited. All three have been based on the comparison of monozygous and dizygous twins, and the three therefore all suffer from the usual methodological deficiencies of twin studies, from the standpoint of genetic analysis. Briefly, there are two major deficiencies for which we must allow. Firstly, there is no way of disentangling the additive and dominance genetic variances, if the comparisons are limited to MZ and DZ twins. Secondly, there is the implicit assumption that environmental factors which make twins similar or dissimilar are equal for both MZs and DZs. On both accounts, the use of twins usually leads to inflated estimates of genetic components. While split MZ twins may be extremely efficient for estimating environmental factors, the limitations of twins for genetics analyses are severe. However, this does not affect the present discussion much, since any biases may be expected to apply equally to all the PMAs. It is then still valid to compare them on a relative basis for evidence of genetic determination.

The first study, by Blewett (1954), gives "heritability" estimates for five of the six PMAs; as noted above, these estimates do not correspond exactly to heritability in the usual sense. In the same year, Strandskov (1954) reported similar data, this time indicating the strength of hereditary influences by a chi-square, testing the difference between MZ and DZ twins. A high χ^2 means that genetic factors are important. Vandenberg (1965) prefers to present similar data as an F-ratio (for 37 and 45 degrees of freedom) of DZ over MZ within-pair variances. The major findings from the three studies are summarized in Table I.

TABLE I

Inheritnace of Primary Mental Abilities

PMA factor	Blewett (1954) h^2	Strandskov (1954) χ^2	Vandenberg (1965) F
Number	0.07	2.12	2.58
Verbal	0.68	12.93	2.65
Space	0.51	14.38	2.42
Word fluency	0.64	5.25	2.48
Reasoning	0.64	0.40	1.40
Memory	(not given)	4.09	1.23

There are obvious discrepancies between some of the results from these studies. Taking the areas of agreement first, all three studies found genetic influences to be important with respect to the verbal, space and word fluency factors, suggesting that none of these had been the subject of strong natural selection in the past. While we can readily accept that verbal and word fluency abilities may not have affected man's capacity to survive and reproduce, at least over much of his history, the neutrality of the space factor is intuitively less obvious. Perhaps the spatial cues employed by hunting man, for instance, are not well represented by the orientation in space task of the PMA test. Vandenberg disagrees with the other two authors on the genetic control of the number factor. If we were content with a majority vote, then we should conclude that numerical abilities have been subjected to fairly

strong natural selection. This perhaps is unexpected, as a "one-two-many" kind of counting would probably have sufficed the majority of mankind until very recently. A possible explanation is that man has been selected for some ability of abstraction that today corresponds to the skills required for a numerical test. We probably know too little yet even to speculate what kind of ability, in detail, that must have been. Blewett is the odd-man-out on the reasoning factor, for which neither Strandskov nor Vandenberg could find any evidence of hereditary control. If we were again prepared to disregard the minority report, this suggests that the ability to recognize sequences and patterns (which is a main feature of the "reasoning" test) may well have been an important factor in survival. It would then have been selected strongly, with the consequent exhaustion of additive genetic variance in the trait. Similarly, the evidence for hereditary control of memory is either weak (Strandskov) or non-existent (Vandenberg). Again, we can conjure up plausible reasons why this may be so.

Obviously, more evidence is required before we can deduce much about the evolution of human intelligence. But enough has been quoted to suggest that perhaps already, we can begin to speculate what kind of "intelligence" was the most important to mankind when it was evolving.

Conclusions

I want to stress again that many statements in the preceding section may have to be revised, or even withdrawn, as more and better evidence becomes available. None of the suggestions that a certain kind of behavior had been important in evolution while another was less so should be regarded as anything more than tentative. My intention was to *illustrate* how knowledge of the genetic architecture of a character may be employed to elucidate the evolutionary significance of that character. For this purpose, data were accepted uncritically and applied as if they were adequate. If the approach suggested is to find any application in the future, then this task should be executed by people competent to judge the data from a behavioral point of view.

Even with fragile evidence, however, the method holds some promise. Some of the suggestions, if substantiated, would throw considerable light on the evolution of behavior. Thus, rats may not have been selected strongly in nature for the kind of ability required to learn a maze, except when they are swimming in water. Individual activity in *Drosophila* may have been more important during evolution than group activity. Excessively or insufficiently reactive rats may not have made good mothers. Memory and reasoning abilities may have been important to man, while verbal skills featured hardly at all in his survival during evolution. Should it become possible to make any such statements with conviction, we shall know more about the evolution of behavior, and thereby be able to interpret laboratory findings more accurately.

It has also been suggested several times in this paper that some particular behavior has been selected for an intermediate optimum. Should this be substantiated, then it is probable that complete genetic studies of behavior will require much attention to genetic correlations. From this point of view, and also to provide good estimates of genetic parameters, studies of the future in behavioral genetics will demand a great deal of genetical sophistication, to match the elegance of the automated measuring techniques and behavioral

competence now widely employed. The further pursuit of some questions suggested in this paper is wholly dependent upon the union of genetical and behavioral concepts, with experimental skills drawn from each discipline.

Summary

Natural selection acting on characters closely connected to fitness is expected to exhaust most of the additive genetic variance in such a character, any remaining genetic variance being largely non-additive. Conversely, characters displaying mostly additive genetic variance, but little non-additive, seem to be those further removed from the organism's natural fitness. The genetic architecture of a character therefore provides some guide to its evolutionary history.

It is proposed in this paper that genetic information of this kind should be employed in an attempt to identify the evolutionary role of behavior more accurately. The genetic basis of the method is briefly reviewed, and the potential use of the method is illustrated with some behavioral data.

References

Agar, W. E., Drummond, F. H., and Tiegs, O. W. 1948. Third report on a test of McDougall's Lamarkian experiment on the training of rats. J. exp. Biol. **25**: 103-122.

Bignami, G. 1965. Selection for high rates and low rates of avoidance conditioning in the rat. Anim. Behav. **13**: 221-227.

Blewett, D. B. 1954. An experimental study of the inheritance of intelligence. J. ment. Sci. **100**: 922-933.

Broadhurst, P. L. 1960. Experiments in psychogenetics. In *Experiments in Personality*, H. J. Eysenck (ed.) London: Routledge and Kegan Paul.

Brody, E. G. 1950. A note on the genetic basis of spontaneous activity in the albino rat. J. comp. physiol. Psychol. **43**: 281-288.

Connolly, K. 1966. Locomotor activity in *Drosophila*. II. Selection for active and inactive strains. Anim. Behav. **14**: 444-449.

Dawson, W. M. 1932. Inheritance of wildness and tameness in mice. Genetics **17**: 296-326.

Dobzhansky, T., and Spassky, B. 1962. Selection for geotaxis in monomorphic and polymorphic populations of *Drosophila pseudoobscura*. Proc. Nat. Acad. Sci. **48**: 1704-1712.

Erlenmeyer-Kimling, L., Hirsch, J., and Weiss, J. M. 1962. Studies in experimental behavior genetics III. Selection and hybridization analyses of individual differences in the sign of geotaxis. J. Comp. physiol. Psychol. **55**: 722-731.

Ewing, A. 1963. Attempts to select for spontaneous activity in *Drosophila melanogaster*. Anim. Behav. **11**: 369-378.

Falconer, D. S. 1960. *Introduction to quantitative genetics*. Edinburgh: Oliver and Boyd.

Falconer, D. S. 1966. Genetic consequences of selection pressure. In *Genetic and environmental factors in human ability*. J. E. Meade and A. S. Parkes (eds.) Edinburgh: Oliver and Boyd.

Fehmi, L., and McGaugh, J. L. 1961. Discrimination learning by descendants of Tryon maze bright and maze dull strains. Psychol. Rep. **8**: 122.

Fisher, R. A. 1930. *The genetical theory of natural selection*. Oxford: University Press.

Hadler, N. M. 1964a. Genetic influence on phototaxis in *Drosophila melanogaster*. Biol. Bull. **126**: 264-273.

Hadler, N. M. 1964b. Heritability and phototaxis in *Drosophila melanogaster*. Genetics **50**: 1209-1217.

Hall, C. S. 1938. The inheritance of emotionality. Sigma Xi Quart. **26**: 17-27.

Hall, C. S. 1951. The genetics of behavior. In *Handbook of Experimental Psychology*, S. S. Stevens (ed.). New York: Wiley.

Heron, W. T. 1935. The inheritance of maze learning ability in rats. J. comp. Psychol. **19**: 77-89.

Heron, W. T. 1941. The inheritance of brightness and dullness in maze learning ability in the rat. J. genet. Psychol. **59**: 41-49.

Hirsch, J., and Boudreau, J. C. 1958. Studies in experimental behavior genetics I. The heritability of phototaxis in a population of *Drosophila melanogaster*. J. comp. physiol. Psychol. **51**: 647-651.

Hostetter, R. C., and Hirsch, J. 1967. Genetic analysis of geotaxis in *Drosophila melanogaster*: Complementation between forward and reverse selection lines. J. comp. Physiol. psychol. **63**: 66-70.

Huntly, R. M. C. 1966. Heritability of intelligence. In *Genetic and environmental factors in human ability*. J. E. Meade and A. S. Parkes (eds.) Edinburgh: Oliver and Boyd.

Huxley, J. 1963. The future of man – evolutionary aspect. In *Man and his future*, G. Wolstenholme (ed.) London: Churchill.

Krechevsky, I. 1932. Hypotheses vs. "chance" in the presolution period in sensory discrimination learning. Univ. Calif. Publ. Psychol. **6**: 27-44.

Krechevsky, I. 1933. The hereditary nature of "hypotheses." J. comp. Psychol. **16**: 99-116.

Kuppusawny, B. 1947. Laws of heredity in relation to general mental ability. J. genet. Psychol. **36**: 29-43.

Lerner, I. M. 1954. *Genetic homeostasis*. New York: Wiley.

McClearn, G. E. (in press). Social implications of behavioral genetics. In *Biology and Behavior: Genetics*, D. Glass (ed.) New York: Rockefeller University Press.

McDougall, W. 1938. Fourth report on a Lamarkian experiment. Brit. J. Psychol. **28**: 328-345.

McGaugh, J. L., Jennings, R. D., and Thompson, C. W. 1962. Effect of distribution of practice on the maze learning of descendants of the Tryon maze bright and maze dull strains. Psychol. Rep. **10**: 147-150.

Manning, A. 1961. The effects of artificial selection for mating speed in *Drosophila melanogaster*. Anim. Behav. **9**: 82-92.

Manning, A. 1963. Evolutionary changes and behaviour genetics. In Genetics Today, S. J. Geerts (ed.) Vol. III, 807-815. Proc. XI Int. Cong. Genet. Oxford: Pergamon.

Roberts, R C. 1965. Some contributions of the laboratory mouse to animal breeding research. Anim. Breed. Abs. **33**: 339-353.

Robertson, A. 1955. Selection in animals: synthesis. Cold Spr. Harb. Symp. quant. Biol. **20**: 225-229.

Robertson, A. 1956. The effect of selection against extreme deviants based on deviation or on homozygosis. J. Genet. **54**: 236-248.

Rosenzweig, M. R., Krech, D., and Bennett, E. L. 1960. A search for relations between brain chemistry and behavior. Psychol. Bull. **57**: 476-492

Rowland, G. L., and Woods, P. J. 1961. Performance of the Tryon bright and dull strains under two conditions in a multiple T-maze. Canad. J. Psychol. **15**: 20-28.

Rundquist, E. A. 1933. Inheritance of spontaneous activity in rats. J. comp. Psychol. **16**: 415-438.

Searle, L. V. 1949. The organization of hereditary maze-brightness and maze-dullness. Genet. Psychol. Monogr. **39**: 279-325.

Strandskov, H. H. 1954. A twin study pertaining to the genetics of intelligence. Carylogia Suppl., Att. IX Int. Cong. Genet., 811-813.

Thompson, W. R. 1954. The inheritance and development of intelligence. Proc. Assoc. nerv. ment. Dis. **33**: 209-231.

Thurstone, L. L. 1941. *The primary mental abilities tests*. Chicago: Science Research Associates.

Tolman, E. C. 1924. The inheritance of maze learning in rats. J. comp. Psychol. **4**: 1-18.

Tryon, R. C. 1929. The genetics of learning ability in rats: preliminary report. Univ. Calif. Publ. Psychol. **4**: 71-89.

Tryon, R. C. 1940. Genetic differences in maze-learning ability in rats. 39th Yearbook Nat. Soc. Stud. Educ. (Part 1) Bloomington, Ill.: Public School Pub. Co., 111-119.

Tryon, R. C. 1942. Individual differences. In *Comparative Psychology*, F. A. Moss (ed.) New York: Prentice-Hall.

Vandenberg, S. G. 1965. Multivariate analysis of twin differences. In *Methods and goals in human behavior genetics*, S. G. Vandenberg (ed.) New York: Academic Press.

Washburn, S. L. 1960. Tools and human evolution. Sci. Amer., Vol. 203 (3), Sept., 1960.

Wherry, R J. 1941. Determination of the specific components of maze-ability for Tryon's bright and dull rats by means of factorial analysis. J. comp. Psychol. **32**: 237-252.

PAPER 27.

Side effects of selection for growth in laboratory animals.

Livestock Prod. Sci. 6, 93-104. 1979

by

R.C. ROBERTS.

Livestock Production Science, 6 (1979) 93—104

SIDE EFFECTS OF SELECTION FOR GROWTH IN LABORATORY ANIMALS

R.C. ROBERTS

Agricultural Research Council Unit of Animal Genetics, Institute of Animal Genetics, West Mains Road, Edinburgh EH9 3JN (Great Britain)

(Received 24 January 1978)

ABSTRACT

Roberts, R.C., 1979. Side effects of selection for growth in laboratory animals. *Livest. Prod. Sci.*, 6: 93—104.

Correlated responses are examined from a biological rather than from a biometrical point of view. Selection for increased body weight in laboratory animals usually, though not always, leads to increases in food intake, gross efficiency and fat deposition, while some aspects of fertility are usually impaired. Increases in fatness are more pronounced in older animals, possibly because selection for early growth requires a high food intake which is not correspondingly reduced as the accretion of lean tissue slows down. Further, fat deposition may be an alternative to heat output, which explains why increases in both fat deposition and in gross efficiency may not be incompatible. It is concluded that many of the adverse side effects of selection for growth are the physiological consequences of increased fatness. In terms of applications to domestic livestock, it is suggested that the undesirable side effects should be controlled managementally by restricting food intake, on the grounds that the simultaneous avoidance of deleterious effects would unduly impair the efficiency of selection for increased growth.

INTRODUCTION

> "Hairless dogs have imperfect teeth; long-haired and coarse-haired animals are apt to have, as is asserted, long or many horns; pigeons with feathered feet have skin between their outer toes; pigeons with short beaks have small feet, and those with long beaks large feet. Hence if man goes on selecting, and thus augmenting, any peculiarity, he will almost certainly modify unintentionally other parts of the structure, owing to the mysterious laws of correlation".

Thus wrote Charles Darwin in "The Origin of Species", and though perhaps the laws of correlation are less mysterious than they were in 1859, some aspects of the phenomenon identified by Darwin are still imperfectly understood. I am not sure whether his examples have all withstood the test of time, and his use of the word "peculiarity" could be debated. But the accuracy of Darwin's prognostication is striking, made as it was before genetics or statistics, as we know them today, had been developed.

Though correlated responses to selection are a common experience, their

genetic interpretation is not always easy. Selection experiments are often done with rather small populations, and the side effects may be more attributable to drift or to inbreeding than to the process of selection itself. Nor is a replication of the study an adequate safeguard against false assignation of cause, for replicates of the same size and mating system will have the same inbreeding; replicates will also have similar gene frequencies and ensuing fixation patterns, perpetuating peculiarities of the gene content of the base population. But in a wider context, inbreeding and drift are mere local complications. Other workers, with different base populations, and where the scale of the experiment or the mating system may also be different, will establish the generality or otherwise of particular outcomes. In an event where drift or inbreeding could adversely affect a selection response, the answer is to increase the effective size of the population at the expense, if need be, of the experimental nicety of replication.

Linkage is another source of correlated responses. Where we have linkage disequilibrium, the susceptibility to it is greatest when crossbred material is selected. In such cases, not even the direction of the correlated response can be predicted, for this will depend on whether the majority of the relevant genes are linked in the repulsion or the coupling phase. But in a special case, the effect of linkage becomes predictable — though it is then essentially indistinguishable from the effect of inbreeding. This occurs when loci affecting the correlated trait (e.g. fitness) display heterozygote advantage. Selection, with the consequent changes in gene frequencies, will generate disequilibrium at loci linked to those under selection; if this reduces the frequencies of favoured heterozygotes, an adverse effect on the correlated trait is predictable, and to the extent that selection increases the homozygosity of chromosomal segments, the effect through inbreeding is identical. If heterozygous advantage is a feature of some genes affecting fitness, then the reduced fitness of selected lines becomes inevitable. Nor do we exclusively need heterozygous advantage — any model specifying the "genetic equilibrium" of the base population will do, though some form of non-additivity is implied. These considerations were discussed in much more detail by Latter and Robertson (1962), and they describe some *Drosophila* selection experiments specifically designed to elucidate the issues. While their selected lines generally did show a reduced fitness (on their measure), the results were by no means uniform, and the reduction in fitness was not always statistically significant.

However, despite the idiosyncrasies of particular experiments, certain generalities emerge over time and space. To take an obvious example, laboratory mice have been selected for increased growth in many places at different times. In general, mice selected for large size tend to become fat as they grow older, they ovulate more eggs — though perhaps being more prone to sterility — and they display a much more phlegmatic behaviour than unselected mice. Typically, at least, the changes are all patently obvious without refined measurement, and we observe the phenomena as side effects of genes that contribute to growth. As we know from formal genetics, pleiotropy is universal, and in

quantitative genetics, it is the source of "true" genetic correlations, as distinct from spurious or fortuitous associations that may arise from linkage, for instance. But accepting all this, it is perhaps strange that these pleiotropic correlations should lead to a rather uniform set of observed correlated responses. Sheer biometrical considerations apart, we must presumably allow for different loci to be segregating in different populations, and for different alleles at the loci that do segregate. With a diversity of gene product feeding into a vast biochemical network, the identity of the phenotypic endproducts, at numerous outlets, demands a remarkably homeostatic developmental system. And yet we know that one mutant gene can create havoc within that same system.

While this stability of outcome poses a formal problem, it may in fact be entirely artefactual. We should distinguish between the direct effects of the genes leading to, shall we say, large size, and effects stemming from the physiological consequences of being large. Thus, if a gene contributes to large size, it will presumably make the ovary bigger and the anterior pituitary gland likewise. More hormone is produced, and with a bigger target organ to absorb it, more eggs are shed as a correlated response. Compare this with a gene that makes the mouse eat more and become fat, the fatness being itself part of the increased size, but also causing sterility. If a mouse becomes too fat to breed, it does not seem quite right to call this a pleiotropic effect of the genes that make it large; if we restrict its food intake to keep it slim, then irrespective of its genes for large size, it will breed normally. This serves to remind us that pleiotropy is not a fundamental property of genes, but rather a technical term that we use to rationalise certain features of the phenotype.

Against this background, we shall now examine a few of the better known side effects of selection for growth in laboratory animals.

GROWTH AND BODY COMPOSITION

There is some confusion in literature on the effect of selection for growth on carcass composition. The statement made earlier that large mice get fat needs to be explained further. The best explanation derives from the work of Clarke (1969), but unfortunately the account of this work is not easily available. Clarke describes his carcass analyses of Falconer's (1973) replicated selection lines for large and small body weight, and one of his main conclusions may be paraphrased thus: when you select for growth up to a certain age, there is little effect on fatness up to the age of selection, but at later ages, large animals become progressively fatter. Taking Clarke's finding as an empirical observation for now, it explains some mysteries. For instance, Fowler (1958) found her large mice to be fat and suggested, quite reasonably, that if animals were selected before they began to lay down fat, any increases in gain would be of lean tissue. Hull (1960) submitted this idea to experimental test by selecting mice for high weight in separate lines at 3, 4½ and 6 weeks of age. Contrary to his expectation, his 3-week line was by far the fattest. But Hull

had measured fatness in all cases at 6 weeks of age; according to Clarke, lines selected for gain before 6 weeks should be fatter by 6 weeks, as Hull indeed found.

These concepts were extended by Hayes and McCarthy (1976), who provided further experimental support for Clarke's generalisation. Their mouse lines had been selected for growth at 5 and 10 weeks of age, respectively. The large line selected at 5 weeks was the fatter by 10 weeks, and fatter still by 21 weeks. Hayes and McCarthy present a model of growth in relation to food intake and efficiency, which they attribute in part to Professor Alan Robertson. The essential features of this model are that in the young growing animal, variation in growth is mostly due to variation in food intake, and animals selected for early growth simply eat more. At later ages, however, as fat accumulates, there is further variation in the partitioning of food between fat and protein. Because, the argument goes, fat is energetically denser than lean, the leaner animals are more efficient, and therefore grow more rapidly. It is these leaner animals that are selected at later ages, whereas animals selected when young are voracious and grow fat.

While the model has the merit of integrating a diversity of experimental results, it becomes more attractive if we can link it to a possible mechanism of appetite control. The literature on this subject is vast, and we cannot even begin to discuss it here. However, some work by Webster and his colleagues (see, for instance, Radcliffe and Webster, 1976) on the Zucker rat, suggests that food intake is closely regulated by the impetus of the animal for protein deposition, while the retention of energy and its storage in lipid may be of little consequence for appetite control. If this is right, animals with a high impetus for growth of lean tissue must have a high food intake, and if they fail to moderate their intake as lean tissue growth asymptotes, then the excess food will be laid down as fat. Nor need such animals be inefficient, for one alternative to laying down fat is to put it out as heat. However, arguments about efficiency became complicated because of the incorporation of water in lean tissue. While further speculation would not be justified, we can perhaps begin to see how growth, food intake, efficiency and body composition may be connected in a way that suggests that the whole complex has to be understood before the different parts can be adequately explained.

A slightly amended formulation of the Hayes and McCarthy/Robertson model is shown in Fig.1. This is rather more speculative than earlier versions, and it is framed largely with the energetic input in mind. An essential consideration is that a large proportion of the energetic input, perhaps as much as 70% of the total, is dissipated as heat, and that this could swamp residual sources of variation. A succint review of the nutritional considerations is provided by Webster (1977). The model aims to depict only the genetical features of the system, summarising the likely sources of genetic variation, as discussed below.

The literature from laboratory animals, on a majority vote, indicates fairly clearly that large animals become fat. Unfortunately, there are not many

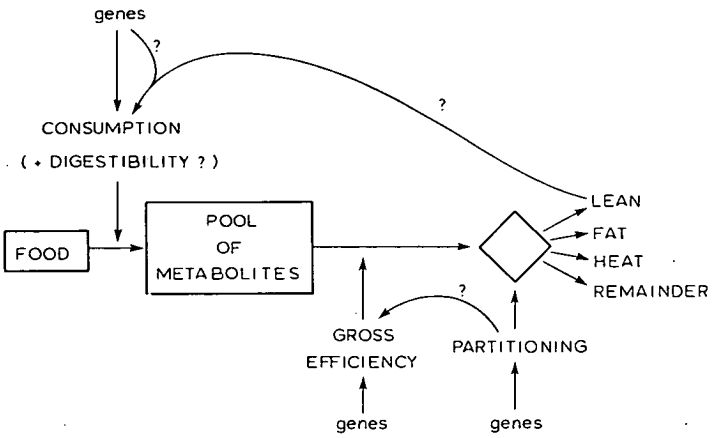


Fig.1. Suggested causal connections between growth, food intake, efficiency and carcass composition. See text for fuller explanation.

studies that have measured body composition over a range of ages. Of those that have, McPhee and Neill's (1976) results with selected mice fit Clarke's prediction very well, while Robinson and Bradford (1969) and Bakker (1974) showed the high lines to be already fatter by the age of selection, and more so at later ages. Of the analyses done at or around the age of selection, increases in the fatness of large mice was reported by Timon et al. (1970), by Sutherland et al. (1974) and by Eisen and Bandy (1977), whereas Lang and Legates (1969) found no increase, and neither did Baker and Chapman (1975), in their case with rats. Eisen et al. (1977), working on the same stock as Lang and Legates (1969), essentially confirmed their results with respect to fatness, but found that another unrelated line selected for large size was already fat by the age of selection. At later ages, Biondini et al. (1968) found substantial increases in the fat content of three lines by 150 days, though one of these had shown little increase over the value of the control line at 112 days.

The age at which the carcass analysis is done affects the interpretation critically. As shown by Sutherland et al. (1974), there are differences among selected large lines in the age at which fat is laid down, which process (to modify Clarke) may therefore be only broadly related to the age of selection. But despite some variation, the general rule seems to be that animals selected for rapid gain become fat, especially as they grow older. The system behaves as if the high food intake necessary to build more protein is not being turned down sufficiently as the protein mass asymptotes.

All the above refers to animals selected under ad libitum feeding, where a connection between rapid gain and fatness was indicated. However, it seems that it may be possible to break this connection. The evidence comes from two separate studies reported by Falconer and Latyszewski (1952) and by Falconer (1960). In both studies, mice had been selected for weight gain under the usual ad libitum system, and also where the food intake was restricted.

The restriction in the first experiment was by feeding less of the usual diet than would normally be consumed. In the second experiment, a low quality diet was specially produced, but fed freely. Both studies showed that when the selected large mice were all reared on a full diet, their capacity for growth had not been markedly affected by selection on different feeding regimes (such differences as were found are not relevant here). But of particular interest, there was substantial evidence that the mice selected on the restricted food intake — in both studies — were less fat even when reared on the full diets, and correspondingly had more lean tissue. Selection under restriction had favoured growth promoting genes which led to less fat, the fat measurements being at the age of selection (6 weeks) in the first study, and at 12 weeks in the second. When grown on the full diet, both lines ate about the same amount of food. Selection on the restricted diet seems to have introduced some additional selection for genes affecting partitioning, in terms of the model presented in Fig.1.

These results clearly suggest the possibility of selecting for growth without excess fat, if that is what is economically desirable. The idea of "breaking" undesirable genetic correlations was pioneered by Cockrem (1959), using body weight and tail length in mouse as an experimental model. Further suggestions of the feasibility of this are provided by the studies of Gall and Kyle (1968), Baker and Cockrem (1970), Rutledge et al. (1973) and, especially, by Eisen and Bandy (1977). However, as Eisen and Bandy rightly emphasize, antagonistic index selection carries a cost in terms of reduced gains, and it may well prove uneconomic to avoid any increase in fatness while selecting for gain.

FOOD CONSUMPTION AND EFFICIENCY OF CONVERSION IN SELECTED LINES

As suggested earlier, it may not be desirable to separate these topics from aspects of body composition, and it did indeed at times prove impossible to omit them from the previous discussion. Generally, lean tissue increases are accompanied by more fat. In a mechanical sense, at least, fatness arises only from hyperphagia, and we should therefore examine food consumption in relation to the selection responses. It is only fair to add that Radcliffe and Webster (1976), quoted earlier, would regard such a mechanical interpretation of fatness as an over-simplification which avoids the issue of appetite control. Nevertheless, it is perfectly reasonable to ask the genetic question: can we explain all selection responses for growth in terms of correlated changes in food consumption, or do we need changes in gross efficiency as well? We could in principle proceed from there to enquire which nutritional aspect — perhaps in physiological or biochemical terms — had been most affected by changes in growth. However, we should also be increasingly subjected to a difficulty discussed earlier, namely that of distinguishing primary or causative changes from those consequential on the changes in size, and therefore in metabolism. An incisive analysis of some of these difficulties, even with single gene substitutions, was provided by Bulfield (1972).

The literature on the food intake and gross efficiency of selected lines of mice show a relatively uniform pattern: large lines generally consume more food (it would be astonishing if they did not) and they also tend to convert it more efficiently. Fowler (1962), Rahnefeld et al. (1965). Lang and Legates (1969), Timon and Eisen (1970), Sutherland et al. (1970), Eisen et al. (1977) and Eisen and Bandy (1977) all agree that the changes are in this direction, and usually significantly so. In other words, there is experimental evidence for some of the sources of genetic variation that are critical for the model presented earlier.

The general pattern seems to be that selection for growth leads to increased food consumption, which in turn increases the gross efficiency of the animal, at least up to a certain level of feeding. Thus, if it takes x units of food just for maintenance, then the animal that eats $(x+2)$ units is bound to be more efficient than the animal that eats $(x+1)$. The genetic interpretation of changes in gross efficiency, stemming from the covariance of a ratio with its own numerator (gain/food with gain), is considered by Timon and Eisen (1970), who conclude that in their material, the increased growth rate resulted mostly from increased food intake, from which followed the increase in gross efficiency. They failed to find any change in net energetic efficiency, confirming Fowler's (1962) report.

Arising from these considerations, it should be noted that across strains, though not necessarily within strains, the more efficient animals are also the fatter ones. Eisen and Bandy (1977) report a positive correlation of 0.90 ± 0.08 between percent fat and efficiency, across replicates. This perhaps is not entirely surprising. The frequent reference to the energetic density of fat — with the implication that it is expensive to produce — is based on its combustible energy. The relevant energy is the one required to synthesize fat, which is much the same as for an equivalent weight of protein (Webster, 1977), before the addition of water gives lean tissue its advantage in efficiency. But the heat exchanges in various reactions, quite apart from any differential loss of heat, does not necessarily mean that it is inefficient to lay down fat, and this apart, fat is usually laid down at an age when the animal's measured efficiency is reducing to zero anyway, as growth slows down and maintenance costs increase. In this situation, it is more efficient to lay down fat than to lay down nothing at all.

The paper by Sutherland et al. (1970), quoted above, reports an experiment of elegant design. It investigates the relationship between growth, food intake and efficiency by selecting for the three separately but contemporaneously from the same base, recording the correlated responses in the two traits not selected directly. All three traits responded, with some unexpected results. The responses in efficiency and in food intake were greatest in the two lines selected, respectively, for those traits. But weight gain was greatest in the line selected for efficiency. This result could not have been predicted, and its interpretation demands caution, in view of Hill's (1971) treatment of drift variance in selected lines. The outcome no doubt reflects the very high genetic

correlation found by the authors between gain and efficiency, both from a component analysis and as a realised genetic correlation. They also provide heritability estimates. With understandably large standard errors attached to such estimates, arithmetic becomes futile, but even so, it does nothing to suggest the generality of the finding that weight gain is greater when selecting for efficiency than when selecting for gain itself.

At the moment therefore, there seems to be little doubt about the concomitant increase in food intake as growth rate is increased. Changes in gross efficiency are also readily discernible, but their interpretation is more ambiguous. From the farmer's perspective, gross efficiency is probably what matters most. From a practical point of view, therefore, one approach might be simply to maximize the weight gain, if this also increases gross efficiency. Any undesirable effects on carcass quality could be controlled by restricted feeding, which under reasonable management, is probably being practised anyway. For direct work on efficiency, measurements of food intake are necessary. This is expensive, because it demands either a high labour input or else sophisticated equipment. In either event, its practical value is diminished.

We should add the obvious caution that results with laboratory animals do not necessarily apply directly to domestic livestock, especially ruminants. An added difficulty is that laboratory animals are usually examined at a given age, whereas the market demands of the livestock industry are concerned more with the fat content, which may be related more directly to weight. Nevertheless, the genetic control of the variables may have much in common for all species, and laboratory animals can help to focus more clearly on the genetic problems in animal breeding.

EFFECTS ON FERTILITY OF SELECTION FOR GROWTH

The fertility of lines of mice selected for body weight has been a subject for comment since the first such experiments. However, MacArthur (1944) and Goodale (1938, 1953) referred to different experiences; MacArthur found that large mice ovulated more eggs and had larger litters, while Goodale had found sterility to be fairly common. Actually, both pioneers have had their experiences amply confirmed. Sterility is common, but of the mice that do give birth to litters, litter size is greater in large mice. Increases in litter size among lines selected for growth have been reported by Rahnefeld et al. (1966), Wilson et al. (1971), Eisen et al. (1973), Falconer (1973), McLellan and Frahm (1973) and by Hanrahan and Eisen (1974). Bradford (1971), on the other hand, observed no changes in litter size, but he did find an increase in sterility, as also did Eisen et al. (1973) and Falconer (1973).

The interpretation of changes in fertility in response to selection for weight is not straightforward, for a variety of effects must be considered. First, any aspect of fertility is susceptible to inbreeding depression. Second, there are complicated relationships between body weight and litter size through maternal effects, as discussed by Falconer (1955) and by Hanrahan and Eisen (1974).

Third, there are further reproductive difficulties that seem to be directly related to excessive fatness.

The first detailed examination of the fertility of mice selected for body weight was that by Fowler and Edwards (1960). Among other effects, they found that sterility in one large strain was due to low libido in the males, while female fertility was unimpaired. Though this conclusion was very clear from their data, common experience since that time has not confirmed the generality of the finding (see later). Fowler and Edwards (1960) showed also that the reduction of fertility in mice selected for small size was in some measure due to hypo-functioning of the anterior pituitary in females.

The effect of inbreeding on fertility is very clearly demonstrated in Falconer's (1973) paper, for litter size first increases and then declines, even in his control lines. Falconer claims that the decline agrees with expectation based solely on previous estimates of inbreeding depression in litter size. The number of productive matings also fell, both in large and small selected lines. Eisen et al. (1973) likewise identify the specific effect of inbreeding on several components of fitness in their selected lines.

Fertility can decline to the extent where the maintenance of selected lines becomes difficult (Roberts, 1967; Eisen et al., 1973). One basic problem is that females do not breed when they become too fat. The length of their reproductive life becomes drastically reduced and their lifetime productivity decreases (Roberts, 1961). It seems clear that natural selection operates against either very large or very small mice, and the improvement in fertility when selection is relaxed or reversed is readily detectable (Roberts, 1966), even in the absence of genetic variation in body weight. As the problem, at least the immediate problem, is an environmental one, the environmental solution of early mating, before fat had accumulated, worked well in one instance (Roberts, 1974). But a better solution with farm animals is the one routinely adopted of limiting their food intake and preventing them becoming fat. The reasons why fat animals, females especially, do not breed are not clear. The cause could be partly mechanical, but the metabolic disturbances arising from obesity may have more profound effects on the steroid sex hormones.

Although the various relationships between body size and fertility have been widely documented, there are aspects which are not well understood. Roberts (1967) reported on crosses between four selected large lines; without any alteration in body weight, the litter size of two-line crosses, used as mothers, was 50% above the mean level of the pure line mothers. In the absence of carcass analyses, it is difficult to be dogmatic, but it does appear that large size per se is not necessarily an impediment to a good reproductive performance.

CONCLUSION

This paper reviews the side effects of selection from a biological rather than from a biometrical viewpoint. I have limited myself mostly to selection for weight gain in the laboratory mouse, partly because the effects in that area

have been well documented and partly because the concepts involved are of applied interest. The side effects are mostly deleterious, through increased fatness and reduced fertility. The details may vary, but generally these two consequences occur. On the credit side, gross efficiency is improved. Referring back to the structure outlined in the Introduction, what we seem to be observing are mostly the physiological consequences of making an animal genetically large. This category of correlated responses can perhaps be termed "pleiotropy once removed", and because of their immediate environmental causation, they are amenable to environmental control by restricting the food intake. It is an old and respectable platitude that breeding and management must go together. The conclusion from this review is that the genetic improvement of rate of gain need not be accompanied by undesirable effects, given adequate environmental control over the improved stock. It would grossly impair the efficiency of a selection programme for increased growth if the undesirable side effects were to be simultaneously avoided.

REFERENCES

Baker, R.L. and Chapman, A.B., 1975. Correlated responses to selection for postweaning gain in the rat. Genetics, 80: 191—203.

Baker, R.L. and Cockrem, F.R.M., 1970. Selection for body weight in the mouse at three temperatures and the correlated responses in the tail length. Genetics, 65: 505—523.

Bakker, H., 1974. Effect of selection for relative growth rate and body weight of mice on rate, consumption and efficiency of growth. Meded. Landbouwhogesch. Wageningen, 74-8, 95 pp.

Biondini, P.E., Sutherland, T.M. and Haverland, L.H., 1968. Body composition of mice selected for rapid growth rate. J. Anim. Sci., 27: 5—12.

Bradford, G.E., 1971. Growth and reproduction in mice selected for rapid body weight gain. Genetics, 69: 499—512.

Bulfield, G., 1972. Genetic control of metabolism: enzyme studies of the *obese* and *adipose* mutants in the mouse. Genet. Res., 20: 51—64.

Clarke, J.N., 1969. Studies on the genetic control of growth in mice. Ph.D. thesis, University of Edinburgh.

Cockrem, F.R.M., 1959. Selection for relationships opposite to those predicted by genetic correlation between two traits in the house mouse (*Mus musculus*). Nature (London), 183: 342—343.

Eisen, E.J. and Bandy, T., 1977. Correlated responses in growth and body composition of replicated single-trait and index selected lines of mice. Theor. Appl. Genet., 49: 133—144.

Eisen, E.J., Bakker, H. and Nagai, J., 1977. Body composition and energetic efficiency in two lines of mice selected for rapid growth rate and their F_1 crosses. Theor. Appl. Genet., 49: 21—34.

Eisen, E.J., Hanrahan, J.P. and Legates, J.E., 1973. Effects of population size and selection intensity on correlated responses to selection for postweaning gain in mice. Genetics, 74: 157—170.

Falconer, D.S., 1955. Patterns of response in selection experiments with mice. Cold Spring Harbor Symp. Quant. Biol., 20: 178—196.

Falconer, D.S., 1960. Selection of mice for growth on high and low planes of nutrition. Genet. Res., 1: 91—113.

Falconer, D.S., 1973. Replicated selection for body weight in mice Genet. Res., 22: 291—321.

Falconer, D.S. and Latyszewski, M., 1952. The environment in relation to selection for size in mice. J. Genet., 51: 67—80.

Fowler, R.E., 1958. The growth and carcass composition of strains of mice selected for large and small body size. J. Agric. Sci., 51: 137—148.
Fowler, R.E., 1962. The efficiency of food utilization, digestibility of foodstuffs and energy expenditure of mice selected for large or small body size. Genet. Res., 3: 51—68.
Fowler, R.E. and Edwards, R.G., 1960. The fertility of mice selected for large or small body size. Genet. Res., 1: 393—407.
Gall, G.A.E. and Kyle, W.H., 1968. Growth of the laboratory mouse. Theor. Appl. Genet., 38: 304—308.
Goodale, H.D., 1938. A study of the inheritance of body weight in the albino mouse by selection. J. Hered., 29: 101—112.
Goodale, H.D., 1953. Third progress report on breeding larger mice. J. Genet., 51: 580.
Hanrahan, J.P. and Eisen, E.J., 1974. Genetic variation in litter size and 12-day weight in mice and their relationships with post-weaning growth. Anim. Prod., 19: 13—23.
Hayes, J.F. and McCarthy, J.C., 1976. The effects of selection at different ages for high and low body weight on the pattern of fat deposition in mice. Genet. Res., 27: 389—403.
Hill, W.G., 1971. Design and efficiency of selection experiments for estimating genetic parameters. Biometrics, 27: 293—311.
Hull, P., 1960. Genetic relations between carcass fat and body weight in mice. J. Agric. Sci., 55: 317—321.
Lang, B.J. and Legates, J.E., 1969. Rate, composition and efficiency of growth in mice selected for large and small body weight. Theor. Appl. Genet., 39: 306—314.
Latter, B.D.H. and Robertson, A., 1962. The effects of inbreeding and artificial selection on reproductive fitness. Genet. Res., 3: 110—138.
MacArthur, J.W., 1944. Genetics of body size and related characters. I. Selecting small and large races of the laboratory mouse. Am. Nat., 78: 142—157.
McLellan, C.R. and Frahm, R.R., 1973. Direct and correlated responses to two-way selection for hindleg muscle system weight in mice. J. Anim. Sci., 36: 442—451.
McPhee, C.P. and Neill, A.R., 1976. Changes in the body composition of mice selected for high and low eight week weight. Theor. Appl. Genet., 47: 21—26.
Radcliffe, J.D. and Webster, A.J.F., 1976. Regulation of food intake during growth in fatty and lean female Zucker rats given diets of different protein content. Br. J. Nutr., 36: 457—469.
Rahnefeld, G.W., Comstock, R.E., Boylan, W.J. and Singh, M., 1965. Genetic correlation between growth rate and feed per unit gain in mice. J. Anim. Sci., 24: 1061—1066.
Rahnefeld, G.W., Comstock, R.E., Singh, M. and Napuket, S.R., 1966. Genetic correlation between growth rate and litter size in mice. Genetics, 54: 1423—1429.
Roberts, R.C., 1961. The lifetime growth and reproduction of selected strains of mice. Heredity, 16: 369—381.
Roberts, R.C., 1966. The limits to artificial selection for body weight in the mouse. II. The genetic nature of the limits. Genet. Res., 8: 361—375.
Roberts, R.C., 1967. The limits to artificial selection for body weight in the mouse. III. Selection from crosses between previously selected lines. Genet. Res., 9: 73—85.
Roberts, R.C., 1974. Selection limits in the mouse and their relevance to animal breeding. Proc. 1st World Congr. Genet. Appl. Anim. Prod., Madrid, Spain, 1974, Vol. I, pp. 493—509.
Robinson, D.W. and Bradford, G.E., 1969. Cellular response to selection for rapid growth in mice Growth, 33: 221—229.
Rutledge, J.J., Eisen, E.J. and Legates, J.E., 1973. An experimental evaluation of genetic correlation. Genetics, 75: 709—726.
Sutherland, T.M., Biondini, P.E., Haverland, L.H., Pettus, D. and Owen, W.B, 1970. Selection for rate of gain, appetite and efficiency of feed utilization in mice. J. Anim. Sci., 31: 1049—1057.
Sutherland, T.M., Biondini, P.E. and Ward, G.M., 1974. Selection for growth rate, feed efficiency and body composition in mice. Genetics, 78: 525—540.

Timon, V.M. and Eisen, E.J., 1970. Comparisons of ad libitum and restricted feeding of mice selected and unselected for postweaning gain. I. Growth, feed consumption and feed efficiency. Genetics, 64: 41—57.
Timon, V.M., Eisen, E.J. and Leatherwood, J.M., 1970. Comparisons of ad libitum and restricted feeding of mice selected and unselected for postweaning gain. II. Carcass composition and energetic efficiency, Genetics, 65: 145—155.
Webster, A.J.F., 1977. Selection for leanness and the energetic efficiency of growth in meat animals. Proc. Nutr. Soc., 36: 53—59.
Wilson, S.P., Goodale, H.D., Kyle, W.H. and Godfrey, E.F., 1971. Long term selection for body weight in mice. J. Hered., 62: 228—234.

RESUMÉ

Roberts, R.C., 1979. Effets secondaires de la selection pour la croissance chez les animaux de laboratoire. *Livest. Prod. Sci.*, 6: 93—104 (en anglais).

Les réponses corrélées sont examinées d'un point de vue biologique plutôt que biométrique. La sélection pour l'accroissement du poids vif chez les animaux de laboratoire conduit généralement, mais pas toujours, à une augmentation de la quantité de nourriture ingérée, de l'efficacité alimentaire brute et du dépôt de gras, alors que certains aspects de la fertilité sont généralement détériorés. L'accroissement de l'état d'engraissement est plus prononcé chez les animaux plus âgés, peut-être parce que la sélection pour une croissance précoce demande une forte consommation de nourriture qui ne diminue pas proportionnellement quand le dépôt de tissus maigres se ralentit. De plus le dépôt de gras pourrait se substituer à la production de chaleur, ce qui explique pourquoi un accroissement du dépôt de gras n'est pas incompatible avec un accroissement de l'efficacité alimentaire. On peut conclure que beaucoup des effets secondaires défavorables de la sélection pour la croissance sont les conséquences physiologiques d'un état d'engraissement accru. Pour l'application aux animaux domestiques, il est suggéré que les effets secondaires indésirables devraient être maîtrisés par une réduction de la quantité de nourriture ingérée, parce que l'élimination simultanée des effets désavantageuses pourrait nuire à l'efficacité de la sélection pour une croissance accrue.

KURZFASSUNG

Roberts, R.C., 1979. Unerwünschte Nebenwirkungen der Selektion auf Wachstum bei Labortieren. *Livest. Prod. Sci.*, 6: 93—104 (in Englisch).

Korrelierte Reaktionen werden mehr vom biologischen als vom biometrischen Gesichtspunkt untersucht. Selektion auf höheres Körpergewicht bei Labortieren führt normalerweise, jedoch nicht immer, zu Steigerungen in Futteraufnahme, Gesamteffizienz und Fettablage, während sich einige Aspekte der Fruchtbarkeit gewöhnlich verschlechtern. Verstärkte Verfettung sind bei älteren Tieren ausgeprägter, möglicherweise, weil Selektion auf Fruhwüchsigkeit eine hohe Futteraufnahme erfordert, welche nicht entsprechend verringert wird, wenn sich das Wachstum von magerem Gewebe verlangsamt. Ausserdem kann Fettablagerung eine Alternative zur Wärmeproduktion darstellen. Dies erklärt, warum ein Anstieg sowohl der Fettablagerung als auch der Gesamteffizienz durchaus vereinbar sein kann. Es wird die Schlussfolgerung gezogen, dass viele der unerwünschten Nebenwirkungen der Selektion auf Wachstum die physiologischen Konsequenzen einer steigenden Verfettung sind. Was die Anwendung bei landwirtschaftlichen Nutztieren angeht, so wird empfohlen, dass unerwünschte Nebenwirkungen durch eine Beschränkung in der Futteraufnahme kontrolliert werden sollten, da die gleichzeitige Vermeidung nachteiliger Effekte die Effizienz der Selektion auf erhöhtes Wachstum unnötig beeinträchtigen würde.

PAPER 28.

Genetic influences on growth and fertility

Symp. on Biology of the House Mouse, the Zoological Society of London, Dec. 1979. Symp. Zool. Soc. Lond. No. 47, 231-254. 1981.

by

R.C. ROBERTS.

Symp. zool. Soc. Lond. (1981) No. 47, 231–254

Genetical Influences on Growth and Fertility

R. C. ROBERTS

Institute of Animal Genetics, West Mains Road, Edinburgh, U.K.

SYNOPSIS

Body weight is a highly heritable trait and responds readily to selection. While the total response is fairly predictable, the pattern of the response varies among populations with different genetical parameters, and also because of accidents of gene sampling between generations. It is postulated that much of the genetical variance found among wild populations is a consequence of their chance genetic origin — the founder effect. While there may be a range of variation over which there is little natural selection on body weight, extreme deviants in any direction can be shown to be less fit. It is argued that there could be similar stabilizing selection operating on fertility, but that during evolution, there has been selection for the largest litters that can survive over an average range of conditions. This usually results in considerable mortality, but also allows an increase in fertility should conditions temporarily improve.

INTRODUCTION

There is by now abundant evidence that genes influence both body weight and fertility. Many mutant genes have been isolated which influence both of these traits, sometimes simultaneously. Such mutant genes, however, are to varying degrees pathological in their effects, and because they are seriously deleterious, occur at very low frequencies. They are not usually detectable in populations where they are not deliberately maintained, and as such, are of no immediate consequence in the context of natural populations. But it is equally clear that other genes, which for present purposes can be labelled "normal" alleles, also contribute to the natural variation both in body weight and in fertility. This kind of genetical variation can be studied in the laboratory, where pedigrees can be kept, by the standard techniques of quantitative genetics. The measure of the importance of genetic influence is frequently expressed as the *heritability*. This is a useful parameter, since it governs the extent to which relatives resemble one another (essentially what we mean when we say a trait is "genetic") and it also allows us to predict

responses to selection. If one were to say that the heritability of body weight is in the region of 30% to 40%, and that of fertility traits not more than half of this amount, few people who know the subject would quibble about its generality. For sure, estimates can be produced that lie outside these narrow ranges, and there are reasons for that. But the generality is such that the *cognoscenti* might not even bother to enquire what species was being discussed – it could equally well be laboratory mice or broiler chickens (not fish, though). In short, body weight tends to be highly heritable – where almost half of the natural variation may be due to genes. Fertility traits are much less heritable, some of them being of very low heritability indeed, e.g. conception rates in cows. Whether the same would apply to conception rates in mice is uncertain, but it would be a good guess that it would.

Two points merit brief further consideration – the different heritabilities of different traits, and their applicability across species or populations. The thinking on the first of these goes back to Fisher's (1930) "fundamental theorem of natural selection". The argument, stripped to bare essentials, states that traits close to natural fitness will have been exposed to millennia of natural selection, and any available genetical variation in the trait will have become largely exhausted in the process. Thus, such traits will have little genetical variance (of this kind) left, and they will show low heritabilities as a result. This does not mean, however, that gene products do not influence such traits – of course they do. It is just that "superior" alleles will have been selected and thus tend to become fixed in any particular population; they will be largely the same for all individuals and they will therefore not contribute to variation between individuals. The considerations stemming from Fisher's theorem refer only to the *additive* genetical variance, on which selection acts. It is beyond the scope of this paper to pursue this further (see Roberts, 1967a, for further discussion), but in the fitness traits especially, there remains a substantial amount (usually) of *non-additive* genetical variance, arising from dominance and interactions between loci. This non-additive variance leads to the observed effects of inbreeding, and its complement – heterosis – to which the fitness traits are particularly susceptible. To the extent that fertility traits are closer to natural fitness than body weight is, the differences between them in general levels of heritability become amenable to genetical interpretation, and are a reflection of the evolutionary history of the traits.

If we accept that fertility and body weight each have similar evolutionary histories across different species, we can accept also that those species will be similar in the genetical architecture of

analogous traits. The generality quoted earlier makes some sense. Having said that, a caveat is necessary. Strictly speaking, heritabilities as estimated refer only to the actual population supplying the measurements and under the conditions under which those measurements were taken. Thus, as Monteiro & Falconer (1966) found, the heritability of body weight in the mouse alters according to the age at measurement, and Falconer (1960a) showed that for a given age, the heritability observed could be influenced by the plane of nutrition. Any particular estimate of heritability should thus be considered as the product of a unique set of circumstances – indeed, the same point could be made of any experimental result. Nevertheless, if we took all estimates of heritabilities of body weight in the mouse at, say six to eight weeks of age, and similarly all heritabilities of the number born in first litters, the two sets of estimates would be virtually non-overlapping. The generality that body weight is more heritable than fertility holds – as a generality – and any exceptions should not cause undue concern. Chance and circumstances will see to it that there are exceptions.

In view of this, what might be the relevance of genetical work with laboratory mice to natural populations of wild mice? Speaking strictly, we cannot tell, because we cannot exercise laboratory procedures under non-laboratory conditions. We can, however, hazard a guess: any differences will be due more to a difference in the conditions, rather than to any gross difference in the genetical make-up of laboratory and wild mice. Some reviewers (Roberts, 1965a, b; Eisen, 1974) have discussed the applicability of genetical work with laboratory mice to domestic livestock, and found adequate parallels. In evolutionary terms this is a quantum jump compared to the applicability of mice to other mice, so perhaps we need do no more than exercise the customary care when extrapolating beyond the range of our data.

It is against this background that the genetical investigations of body weight and fertility will be considered.

GENETICAL VARIATION IN BODY WEIGHT

Differences in Body Weight between Strains and Populations

There is abundant evidence that inbred strains of mice, kept under the same laboratory conditions, differ in body weight at the same age. The evidence will not be reviewed in detail here – see Poiley (1972) for extensive data on the topic. This fact alone is sufficient proof of genetical variation in body weight, though it tells us nothing

of the nature of such genetical variation, nor how easily it may be exploited by selection. It tells us nothing either of the adaptive significance of body weight (if any), for the origin of such genetical variation is presumably adventitious, or largely so, and arose from the historical separation of inbred strains during their formation. We may suppose, however, that what pertains to laboratory strains will apply, at least in part, to wild populations that are reproductively isolated. Direct comparisons between natural populations are of course complicated by possible – indeed likely – differences in the environment. Even so, Berry and his colleagues (Berry & Jakobson, 1975; Berry, Peters & Van Aarde, 1978; Berry, Jakobson & Peters, 1978) have argued convincingly for genetical variation among island populations of wild mice, including genetical variation in body weight. These studies ranged from the Faroe Islands to the Australian sub-Antarctic. The Faroe populations were compatible with the colonization of separate islands by small numbers of effective founders, and the analogy with the differentiation of inbred strains (above) is clear. This is not to say that Faroe mice are highly inbred by laboratory standards; it is just that similar forces are at work, and that genetical differentiation could not occur unless there had been initial genetical variation.

Because of this founder effect, we have to be cautious about inferring genetical differences between populations when wild-caught samples are brought into the laboratory for further study. The numbers caught are usually rather small, and of those caught, fewer breed. In addition, who is to say whether mice successfully trapped are a random sample of the population they are purported to represent? With this caveat, however, studies on wild mice in a standard laboratory environment confirm the existence of ample genetical variation, either between or within wild populations, with respect to body weight (Plomin & Manosevitz, 1974; Barnett *et al.*, 1975; Ebert & Hyde, 1976; Lynch, 1977). Of these, only Plomin & Manosevitz compared populations derived from different localities, two from Texas and one from Colorado. Body weights were clearly different at 200 days, though not at 22 days.

Vagaries of sampling apart, differences between populations tell us nothing about the adaptive significance of body weight. It would be equally plausible to argue that each population is carefully adapted to its own ecological niche as it would be to maintain that body weight is of little consequence, reflecting little more than the historical accidents of founder effect with subsequent non-directional drift. Common sense would dictate that perhaps both elements might have contributed to the current state, and only supplementary

evidence can suggest their relative importance. Such supplementary evidence may derive from two distinct sources. The first is well illustrated by Berry, Jakobson & Peters (1978), who calculated genetical distances for several characters among different populations from the Faroes. In their analyses, they included body weight, organ weights, allozyme variants and indices of skeletal shape. What they found was that distances calculated from different characters were poorly correlated with each other. Even correction of the skeletal parameters for body size failed to improve the correlations of distances among the skeletal parameters, suggesting to the authors that size *per se* was unimportant. Had body size lent coherence to the remaining skeletal data, then the adaptive significance of body size would have been indicated. As it is, the conservative conclusion is that the adaptive significance of body weight could not be established, within the bounds of this data set.

The second approach is to examine the genetical properties of mouse populations by selection. The responses themselves offer some guide to the evolutionary history of the trait, as outlined earlier. In addition, the reproductive performance of lines selected for body weight allows the correlated effects on fitness to be examined.

Responses to Selection for Body Weight

Selection for body weight in the mouse has a long history, dating from the pioneering studies of Goodale (1938) and MacArthur (1944). Comprehensive reviews of the field, up to the dates of publication, have been provided by Roberts (1965a) and by Eisen (1974), and specialized aspects of the topic were further discussed by Roberts (1979), bringing the bibliography more or less up to date. To avoid repetition, if nothing else, no attempt will be made to be equally exhaustive here; references will be chosen purely to illustrate the main findings.

In the same way that Berry and his colleagues showed founder effects on island colonies, so do selection experiments reflect the gene content of their various base populations. More than that, however, even a given gene content does not necessarily yield a predictable outcome, because during selection – or even propagation without selection – genes can be lost through accidents of sampling. This may be true of genes even favourable to the direction of selection, particularly if they are at low frequency in small populations. This is the essence of Kimura's (1957) oft-quoted paper on chance fixation, extended by Robertson (1960) in a theoretical treatment of limits to selection. Thus, a selected line may become

fixed for a particular allele even though a better one had originally been available, because the better allele was accidentally lost in the process. It will be intuitively obvious that this accidental loss will be less likely as the population size increases, as the frequency of the favourable allele increases, and as the magnitude of its effect increases. Chance fixation can be a major factor in the limits to selection ultimately reached; initial responses will be subjected to general drift – chance fixation being a special case.

The repeatability of selection responses was tested by Falconer (1973), with specific reference to body weight at six weeks of age in the mouse. He had six lines selected for high body weight, six for low and six unselected control lines. All derived from the same base population. The effects of drift were very clear in the control lines. Initially their mean weights varied from about 22 g to 23.5 g. After 20 generations of random propagation, the range was fully from 21 g to 25 g. The initial sampling, even from the same families for pairs of lines, had generated significant differences in body weight. Subsequently, the control lines diverged further because of drift. The effects of random drift were equally evident (or almost so) in the selected lines. After 20 generations, the large lines varied in mean weight from about 32 g to 35 g, the small lines from 13 g to 16 g. So, despite the differences between replicated lines due to drift, they nevertheless reached similar end points. The overall effect of drift must therefore be judged to be small, compared to the effect of the selection, even with the rather small population sizes. The effective number of parents (selection being within families) was never more than 32 in Falconer's lines, and frequently less because of some sterility.

Whether the responses are judged to be similar or dissimilar is very much a matter of outlook and emphasis. Having pointed out the similarities of the weights reached after 23 generations of selection, Falconer was still able to show "forcibly how dissimilar the replicates were over the first part of the selection", especially when the divergence between upwards and downwards selection was compared. Falconer said that his results led to a "clear warning" that deserves to be quoted in full: "single selection experiments on the scale of one of these replicates can be very misleading about the rate of response, and particularly about the asymmetry, if judged from the first 5 or even 10 generations". In conjunction with this warning. Falconer reported that the range of realized heritabilities among replicates was from 25% to 46% for the high lines, and from 16% to 50% for the low.

Falconer's experiment has been given some space here because it is

the most comprehensive on record. Though the point did not figure among Falconer's declared objectives, we might pause to consider what kind of framework it provides for thinking about populations of mice in the wild. The main difficulty here is the uncertainty about effective population sizes in the wild. Even though Laurie (1946) was able to capture 2368 mice from one wheat rick, the original invasion may have been a small number of effective parents. If we use Falconer's effective number of 30 or so as a not unrealistic model for many populations, then we should not overestimate the effect of drift on body weight in the short run. And despite drift errors being cumulative (Hill, 1971), they need not amount to all that much in the long run either. As the population size increases, drift becomes less; but if the population size decreases, then inbreeding inevitably occurs, and though this may generate considerable genetical differentiation *at the time,* the concomitant loss of genetical variance will allow less scope for further genetical differentiation thereafter. After Falconer's (1973) report, the population size of all experimental lines in Edinburgh was doubled; after a further 40 generations of random mating, the control lines are no more divergent now than they were then. Whether this lack of further divergence is due to the increased population size or to the accumulated effects of inbreeding (i.e. to loss of genetical variance) is a moot point. Generally speaking, unless population sizes are relatively small and relatively stable, drift will not be detectable as a gradual accumulation that proceeds indefinitely. It will occur in fits and starts, corresponding to a bottleneck in population size through some crisis. The small sample may well generate a change in mean body weight, while the inevitable inbreeding will reduce the scope of further drift. Any migration between populations will of course have the opposite effect, and prevent genetic isolates being developed.

The amount of selection practised in Falconer's experiment was presumably far greater than any selection on body weight operative in the wild. We need not suppose, therefore, that any differentiation through drift among wild populations will be swamped by the effects of selection. It may be argued, however, that unlike laboratory populations, wild mice may be subject to spatial and temporal variation in the environment. To the extent that different populations may become adapted to specialized niches, selective forces are inescapably implied. The balance of the evidence is probably against this happening, as far as body weight is concerned, as discussed by Berry and his colleagues, quoted earlier.

Other reports testify to the high heritability of body weight, notably the extensive studies of Eisen and his colleagues in North

Carolina. Eisen (1978) found heritabilities of six-week weight to be from 42% to 55%, depending on the method of estimation. A point of particular interest in the North Carolina work is the simultaneous selection for body weight and traits negatively correlated with it. Large mice generally have larger litters and longer tails, and vice versa, among other correlated traits. Eisen (1978) selected for the antagonistic relationships involving litter size and Eisen & Bandy (1977) selected similarly in a replicated experiment involving body weight and tail length. In agreement with theory, the changes in body weight were less when selection involved also a second trait. The implications for what may happen in the wild are obvious. If body weight is part of an adaptive complex involving other traits (as will undoubtedly be the case), then the scope for changing body size – depsite its high heritability – may be severely restricted if there is concomitant selection for traits negatively correlated with it.

The growth curve of the mouse, as in other mammals, can be described in general terms as sigmoid, and a variety of mathematical functions have been employed to obtain the best statistical fit. The logistic equation usually proves to be as good as any. Eisen (1976) comprehensively reviews the effects of selection on growth curves, with the following general conclusion: although the constants in the equation can be shown to have been altered, the overall *shape* of the growth curve is generally unaffected, the differences being mostly due to the re-scaling of the two axes. A study by McCarthy & Doolittle (1977) set out to change the shape of the curve by a variety of procedures: to change five-week and 10-week weights in opposite directions, or else to change one while holding the other constant. While their attempts were not uniformly successful, and in some cases agreed rather poorly with prediction, they nevertheless showed that it was perfectly feasible to alter the shape of the curve. It is only fair to add, however, that they were most successful with the somewhat sophisticated procedure of restricted indices. As a matter of personal opinion, I should find it hard to imagine selective forces in the wild operating in this manner, as there must be more urgent matters demanding selective attention. But it would be foolhardy to dismiss the possibility, particularly if conditions were to favour early maturity (rapid early growth) combined with a small mature size. Other possible combinations would seem to me to be even less plausible.

To conclude this section, body weight has been shown to be highly heritable, and on the argument presented earlier, this is compatible with the idea (though not proof of it) that body weight is not a major component of natural fitness. It is reasonable to postulate

that the variation in body weight found in the wild is therefore either a consequence of chance genetical origin (founder effect), or else the product of nutrition and other environmental variables. And as we saw, drift may also contribute to some differentiation between populations. It would be a worthwhile experiment, for those with an appetite for handling wild mice in captivity, to select them for body weight, to test whether the genetical parameters for wild mice resemble those of synthetic laboratory populations, and lead to similar responses. My guess is that they would, and if this were so, it would be grist for the mill of those who argue against the adaptive significance of body weight. We should be wary, though, of pushing that argument too far. If we think of the genetical situation within populations, there may be a range of body weights around the mean where all mice are more or less equally fit, though possibly for different reasons. For instance, the larger mice may be more successful in establishing territory, but the smaller mice may more easily meet their nutritional demands. But there must be bounds on permissible departures from a limited range, beyond which extreme size in either direction becomes a crippling handicap. Whether a trait is judged to be adaptive or not is nothing more than a statement of the amount of variation in that trait that we are prepared to accept, and it is merely a question of degree.

In the next section, the influence of body size on fertility – which is more obviously a component of fitness – will be examined, to identify some of the reasons why extreme deviants in body weight may be less fit.

Effects of Body Weight on Fertility

Selection for body weight leads to a well-documented correlated effect on litter size: larger mice have larger litters, while small size leads to small litters (see Eisen, 1974; Roberts, 1979). Other things being equal, large mice should therefore be fitter. Other things, however, are not equal, and large mice have their own reproductive difficulties. Indeed, Lerner (1954) argued that directional selection for any metric trait would lead to reduced fitness, and that part of his argument has never been seriously challenged.

The problem with large mice is their proneness to sterility. It is of no use to them to have potentially large litters if they have no litter at all. Eisen, Hanrahan & Legates (1973) and Falconer (1973) both illustrate the problem in sharp relief. Roberts (1967b) reports a large line that was lost through sterility. We need go no further to

appreciate that there *has* to be some kind of stabilizing selection for body weight. Extreme deviants in either direction are not fit, and that is that.

Some of the reproductive difficulties of large mice seem to arise because they get too fat. An offshoot of the line that was lost (above) was saved by mating it at an earlier age (Roberts, 1974) before fat had accumulated. The fertility of another large line was helped considerably by relaxing selection (Roberts, 1966), even though there was no additive genetical variance in body weight remaining in that line. This example proves that at least some of the negative correlation between large size and fertility is environmental in terms of its immediate origin. However, even though the immediate physiological cause may be environmental, it does not remove the ultimate genetical involvement if large mice inexorably (though perhaps not unavoidably) get fat. Many workers have found that large mice get fat, though in some cases the increased fatness is not apparent until later ages (Robinson & Bradford, 1969; Timon, Eisen & Leatherwood, 1970; Bakker, 1974; Sutherland, Biondini & Ward, 1974; McPhee & Neill, 1976; Hayes & McCarthey, 1976; Eisen & Bandy, 1977; Eisen, Bakker & Nagai, 1977).

What might happen in the wild? If large size leads to fatness and fatness to sterility, one answer would be to advance sexual maturity, so that some breeding is done before the fat accumulates. We may imagine that there must be some selection anyway for early maturity and rapidity of reproduction. But in the case of large mice, this would not seem to be adequate compensation for the drastic shortening of the length of their reproductive life. Roberts (1961) compared the lifetime production of two large strains and two small strains. The two small strains each produced 11 litters over their lifetime, as against three and five, respectively, for the two large strains. The result was that the small mice weaned almost twice as many offspring as the large mice. More striking, however, was the lifetime production of a cross between a large and a small strain; the cross-bred mice weaned three times as many offspring, over their lifetime, as the better of the two parental strains.

It may be only partly relevant that the parental strains of this cross differed in body size. Roberts (1967b) reported improvements in fertility among crosses of large lines, though the lifetime production of those crosses was not examined. But we may speculate in this context, given the superiority of hybrids, whether there may not be some selection in the wild for any propensity to out-cross — to seek a mate from outside the population in which the mouse grew up. This could be regarded as an extension of the incest taboo, and

we can at least specify one of the conditions which might favour it. It is easiest to express the notion verbally if we allow ourselves some "selfish gene" thinking. If a mouse (or any diploid organism) produces a hybrid offspring, then only half of the gametes of that offspring will transmit the gene that we have in mind. If, instead, the mouse mates within its own population, then there is some probability that the gene we are monitoring becomes homozygous with a replica of itself, and when such an offspring in turn breeds, all of its gametes will transmit the gene in question. We can therefore see how the balance might swing: out-crossing will become favourable when hybrid offspring, with single copies of the gene, can transmit it more frequently than the more inbred offspring with two copies. Following an out-cross, there will be less advantage to further out-crossing, and offspring with two copies of the gene may do better than hybrids, until inbreeding depression (see later) reduces fertility again. Yanai & McClearn (1972a, b) were able to show the preference of females for mating with unrelated males, both among inbred and random-bred mice. We therefore have a behavioural mechanism for promoting higher fertility by producing hybrid offspring preferentially. But the selective advantage of this depends on the reproductive superiority of these hybrids, as noted above.

The Mediation of Gene Effects on Growth

We shall now examine briefly by what mechanisms genes may influence body size, and consider the basis of some of the genetical variation we may observe.

Increased growth could be obtained in one of two ways. Either the mouse could eat more food, or else the same amount of food could be used more efficiently. Selection for increased growth in the laboratory has generally altered both: larger mice eat more and also convert it more efficiently. Falconer (1960a), Fowler (1962), Rahnefeld *et al.* (1965), Lang & Legates (1969), Timon & Eisen (1970), Eisen & Bandy (1977) and finally, Eisen, Bakker & Nagai (1977) all agree that both food intake and efficiency are altered by selection. In addition, Sutherland, Biondini, Haverland *et al.* (1970) showed that both food intake and the efficiency of conversion respond to separate selection, confirming that each trait is under some genetical control. There is a complex relationship between intake and efficiency which affects the interpretation of these results. It is inappropriate to seek out the complexities here, but the main feature of the relationship is the following. If an average-sized mouse needs, say, 15 g of mouse food to keep itself alive for a week, without

additional growth, then the mouse that eats 16 g, and grows a bit, is clearly more efficient than the mouse that eats 15 g and whose weight stays still. Against that, a mouse that gets by on 14 g (still without growing) is more efficient than the mouse whose intake of 15 g just keeps it going. But at marginal intakes, there is usually a positive correlation between intake and efficiency. At the other end of the scale, excessive voracity can lead to inefficiency, if the high intake of food is not converted fully into a weight gain. It could be accompanied by a higher heat loss, or an energetically costly body composition, e.g. excess fat.

If laboratory results can be translated into the field (or barn), what kind of selection on appetite and efficiency might we find? It seems reasonable to suppose that high efficiency would always be at a premium, and though there is some genetical variation in the trait, its heritability is generally lower than that of body weight. It is likely, therefore, that over its evolutionary history, there has been considerable selection for efficiency. It is also plausible to argue that because of the complex relationship between appetite and efficiency, outlined above, there may have been some selection for an intermediate level of intake; mice eating less or more may have a lower efficiency. A low efficiency might not matter as long as food is plentiful and constantly available. But mice that let their efficiencies slip in times of plenty leave descendants who are poorly equipped to meet the next shortage, for survival means the surviving of crises. There may therefore be recurring cycles of weeding out of inefficient mice.

Coleman (1978), in a penetrating review, speculates on the kind of selection that occurs in the wild with specific reference to lipid metabolism. He notes that several species of desert rodents, when brought into the laboratory, can develop symptoms of clinical diabetes. Many animals become hyperphagic, obese, hyperinsulinaemic and show some glucose intolerance, while a few might develop the more extreme symptoms of hyperglycaemia, glucosuria and a ketonic form of diabetes. Coleman points out that in the feral state, these rodents have a limited food supply and develop normally. Further, the very features that cause problems in the plenitude of the laboratory are associated with a metabolically thrifty genotype that, Coleman speculates, is the product of natural selection. During periods of excess of food, the potential hyperphagia, coupled with an increased rate of lipogenesis, allows the rapid accumulation of fat stores for use in days of privation. It is pointed out that obese mice may survive up to 30 days of starvation, whereas normal mice would be dead in two or three days. Coleman suggests that hyperinsulinaemia,

as the most consistent feature of the syndrome, may be the key to the improved lipid anabolism. Whatever the exact nature of the mechanisms, we have a model of the kind of selection that may operate on appetite control, with its consequences for improved efficiency.

Although he does not note it as a specific concern, there is a corollary of Coleman's review that we should note *en passant*. It is that genes beneficial in the wild may be highly deleterious in the laboratory, and find themselves selectively eliminated.

Any discussion of the mediation of gene effects on growth at the biochemical level is both beyond my scope and outwith my competence. The reader is referred to Shire's (1976) exhaustive review of genetical variation in endocrine systems as an excellent place to start. Shire documents abundant evidence of genetical variation, both in various aspects of hormone production and in the responses of target organs. As one example of the hormonal changes brought about by selection for body weight, Pidduck & Falconer (1978) found that increased growth in their strains was partly due to an increased amount, or activity, of circulating growth hormone, while reduced growth was due, again in part, to a reduced sensitivity of the target organs. Clearly, there is immense scope for selection at this level, though a caveat is necessary: just because we observe changes in mechanisms following a change in body size, we need not suppose that the changes in mechanisms were necessarily a direct cause of differential growth; they could just as well be the consequences of differential growth. As an example, we may consider cell number and cell size. It is obvious that if a mouse is to be bigger, then it must either have more cells or bigger cells, or some combination of the two. Robinson & Bradford (1969) suggested that selection for body weight alters cell number rather than cell size, a conclusion to which Priestley & Robertson (1973) somewhat cautiously lend support. In contrast, Falconer, Gauld & Roberts (1978a) found that both cellular components had been altered by selection, and that in some organs, changes in cell mass were as great as changes in cell number. More than that, when large and small mice were compared at the same weight (as distinct from the same age), then the organ sizes were the same and the number of cells was the same; from which it follows that cell size was also the same. Given the size of an organ, then its cellular components became predictable irrespective of strain or age. In other words, the effect of the selection had been to alter developmental age relative to chronological age. Large mice simply grew faster, and one of the results was to increase both cell number and cell size. It is therefore impossible to

say that the changes in the cellular components were in any sense the cause of a change in body weight.

Before we dismiss the cellular basis of growth regulation, however, we should note the implications from aggregation chimaeras between large and small mice (Roberts *et al.*, 1976; Falconer, Gauld & Roberts, 1978b). The proportion of large cells in any individual mouse is a matter of chance, and can vary over the whole range. The body weight of the ensuing chimaera is linearly and directly proportional to the number of large cells, as if growth depended on the cellular genotype throughout the whole body. No particular organ acted as if it was controlling growth. The nine organs included in these studies, taken together, accounted for all of the variance in growth; indeed, any one of them on its own gave a reasonable prediction. If there is a growth-controlling organ, then its cellular composition must correlate highly with those included, which in turn correlated highly among themselves.

The purpose of this section has been two-fold. The first was to hint – no more – at the manifold nature of the raw material on which selection for body weight can act. The second was cautionary: to suggest that we ought not to be over-anxious to deduce causation from association. It may be tempting, at first, to link an animal's growth to the food that it ingests, and the efficiency with which it converts its food into animal product. This, however, ignores the problem of what controls the animal's appetite in the first place. Radcliffe & Webster (1976) postulate that food intake is closely related to the animal's impetus (rats, in their case) for laying down protein. Certainly, there is the well-known phenomenon of compensatory growth, following a period of inadequate feeding, whereby animals tend to revert to a normal weight for age. Indeed, were this not to occur under wild conditions, animals could never recover from temporary deprivations. Growth control must therefore ignore ephemeral perturbations; this may be one of the reasons why the growth control system has proved so intractable to experimental attack.

GENETICAL VARIATION IN FERTILITY

Responses to Selection

Selection for fertility traits has not enjoyed the attention expended upon various aspects of body size and growth rate. The first definitive account of a selection experiment for litter size was given

by Falconer (1960b), though he had published a preliminary account earlier (Falconer, 1955). He produced a high line with a mean litter size of nine live young at birth, and a low line with a mean of six. The heritability of the divergence between the two lines was only 13%; the high line on its own gave a low estimate of 8%, though the low line showed a higher value of 23%. However, because the high line was more variable, the actual responses were not so asymmetrical as the heritability values.

Subsequent experiments have in general exceeded Falconer's somewhat modest responses. Bateman (1966) using what was essentially a form of mass selection, generated a two-fold difference (with a high line of 11) over 12 generations. Bateman does not quote heritabilities but he was, at least at times, selecting intensely. Bakker, Wallinga & Politiek (1978), selecting for large litter size only, increased it from eight to 14, over 29 generations, with a realized heritability of 11%. Eisen (1978) reports the impressive value of 16 young born in a line selected for litter size, though he had started from the high base level of 12. The heritability approached 20%, depending on method of estimation. All of these studies give uncomplicated results. However, Bradford (1968, 1971) presents some interesting variations. He increased litter size from nine to 12 over 11 generations, from a cross-bred derived from eight inbred lines. He was less successful when selecting from a four-line cross, his improvement being only about one offspring per litter. He further selected from the four-line cross after it had previously undergone seven generations of selection for weight gain, and this time failed to improve litter size at all. He was equally unsuccessful when selecting for increased litter size following superovulation. Furthermore, Bradford did not observe the usual correlated effect on litter size when selecting for weight gain. He noted that the genetical correlation between the two traits varies among populations. Batten & Berry (1967) had independently come to the conclusion that body size and litter size need not be correlated; indeed, they went a step further, and claimed that in the case of their island mice, natural selection had operated against such a correlation.

The results of laboratory experiments, on balance, indicate that additive variance for litter size is usually present, though in variable amounts, and that the trait usually responds to selection. On the genetical argument presented in the Introduction, this would suggest that litter size is less closely related to natural fitness than is sometimes supposed. Batten & Berry (1967) make the same point, and we shall return to it briefly in a later section.

Non-additive Genetical Variance in Litter Size

Non-additive variance stems from dominance and interaction effects, and is defined as that part of the genetical variance not amenable to selection. It may nevertheless contribute to the resemblance between certain classes of relatives, particularly full sibs. Crudely, non-additive variance refers to the special effects of combinations of genes, either at the same locus or at separate loci. When an animal breeds, these combinations are broken up and we observe only those effects that genes exercise singly, giving rise to the additive variance. Non-additive variance comes into play during inbreeding and crossing, where levels of heterozygosity (among other genetical effects) are altered. It is the basis of inbreeding depression and heterosis.

There is no room here to review the copious literature on inbreeding and crossing in the mouse. Eisen (1974) provides access to this literature. Suffice to say that litter size declines by about half a mouse per 10% increase in inbreeding coefficient, and that litter size is restored on crossing. Even standard laboratory inbred strains usually (though not inevitably) show heterosis in litter size, despite the fact that such strains are the peculiarities that have survived the inbreeding process, and thus represent nothing except themselves. The vast majority of lines fail to withstand inbreeding, and become extinct through infertility and inviability. The survivors are therefore not random representatives of the base population from which they were drawn. Falconer (1971) was able to capitalise on this fact by forming nine inbred lines from his strain selected for high litter size (see earlier). Inbreeding depression in litter size immediately set in at about the expected rate noted above. However, by maintaining sublines and practising selection for litter size, Falconer was able to maintain four of the lines through 11 generations of sib-mating. At that point he crossed them and the derived cross-breds had a mean litter size of 1.5 mice above that of the original selected strain. Falconer postulated that rare recessive genes, perhaps as many as 30 or possibly more, had arrested the original response, and that these recessive genes had been exposed by inbreeding and eliminated by selection.

What do these general considerations lead us to expect with wild mice? An extension of the argument that traits close to fitness display little additive genetical variance is that such traits should also have considerable non-additive variance left. Traits close to natural fitness should therefore be particularly susceptible to the effects of inbreeding. As far as litter size is concerned, there has been little systematic work on the genetical parameters of wild-caught mice

Two reports suggest that wild-caught mice are not particularly sensitive to inbreeding. Lynch (1977) reported that litter size in wild-caught mice declined with inbreeding at the standard laboratory rate, while the mean litter sizes of her surviving sublines altered little over six generations of sib mating. Connor & Bellucci (1979) similarly employed extensive subline replacement, but even so, five of their ten inbred lines, from wild-caught mice, failed to survive the inbreeding. Litter size declined under their conditions, but only slowly at first. They deduced that inbreeding was being counteracted by natural selection, possibly involving heterozygous advantage. They certainly found substantial heterozygosity in four of their five lines, even after 20 generations of sib-mating.

If litter size in wild mice was one of the major determinants of natural fitness, then inbreeding might be expected to have drastic effects on it. The evidence from the two studies just quoted is that this is not so, and that litter size in the wild has not been subjected to previous natural selection much different to that pertaining under laboratory conditions.

The Components of Fertility

Litter size is a complex trait determined sequentially by ovulation, fertilization, implantation and embryonic survival – even without the perinatal and postnatal hazards that determine the number of offspring that themselves survive to breed. We shall ignore here any genetical influences of the male or litter size, and female sterility will also be excluded. We shall concentrate on the normal range of variation found among fertile animals.

Ovulation rate responds to artificial selection. Land & Falconer (1969) selected both for natural and induced ovulation rates, with substantial responses. The lines selected for high and low natural ovulation, however, did not differ in the number of young born, despite a difference of seven ova shed (21 v. 14). Land & Falconer's induced ovulation lines, on the other hand, differed by about two young at birth, when allowed to ovulate naturally. The genetical correlation of 0.33 which they report between natural and induced ovulation shows that, despite some genetical overlap, the two traits are substantially different, as Bradford (1968) had found in his selection programme. Land (1970) was able to show genetical influences both on FSH activity and on ovarian sensitivity, and that both are positively correlated, genetically, with body weight. This explains why selection for body weight often (though not invariably) changes ovulation rate, and vice versa.

The most extensive study of selection for components is that reported by Bradford (1969). He selected separately for ovulation rate and embryo survival, both in the presence and absence of superovulation. Briefly, embryo survival responded to selection in both cases, increasing litter size by two mice or so in the untreated line. Litter size did not increase in the superovulated line, but embryo survival improved, in proportionate terms, because of a reduction in ovulation rate. Selection for high natural ovulation gave a response of about two ova over ten generations, but without any increase in number born, as Land & Falconer (1969) had also found. Selection for a high induced ovulation rate gave no response.

By and large, therefore, genetical manipulation of the components of litter size has little effect on the number born. Nevertheless, litter size itself *can* be manipulated genetically, so what happens to the components? Falconer (1960b, 1963) found that the response in his high line was entirely attributable to increased ovulation rate, while in his low line, ovulation rate had not been altered but embryonic deaths had increased markedly in the post-implantational stage. Bateman (1966), working with similar material, essentially confirmed Falconer's result, except that the embryonic mortality in his low line was distributed evenly before and after implantation. Bradford (1969) also found that his high litter size line had increased in ovulation (without affecting mortality), noting that the reciprocal effect when selecting for ovulation had not been found (see above). Though perhaps the evidence is too meagre to generalize, it is so far entirely consistent. Selection for litter size yields qualitatively different responses in the two directions: high litter size means more eggs, low litter size greater mortality.

Studies comparable to the laboratory ones on the components of litter size were conducted on wild mice by Batten & Berry (1967). Their material derived from several island and mainland populations. Ovulation rates were low by laboratory standards, seldom exceeding ten. Nevertheless, they found extensive embryonic mortality; a fairly constant fraction of one-third of all eggs were lost, more of the losses occurring before implantation than afterwards. The authors invoke deleterious genetical factors to explain these deaths, but this is open to question. Certainly, inbreeding studies on wild mice (Lynch, 1977; Connor & Bellucci, 1979) do not suggest that wild mice are particularly prone to the exposure of recessive lethals. Further, Southwick's data (quoted by Batten & Berry, 1967) show that embryonic death increased with population density. It seems likely that much of the embryonic death in wild mice is of environmental origin, and that the lethality has not been eliminated by natural

selection since it affords a reservoir of higher fertility should environmental circumstances prove favourable.

So, what price litter size in wild populations? As Batten & Berry (1967) point out, optimum litter size is probably submaximal, since survival may be reduced in large litters. And as we saw earlier, mice with large litters may have a shorter length of reproductive life (Roberts, 1961). In that study there was the complication of large body size. But in another case, even a line selected for high litter size (14 as against eight for the control line) was overtaken by the control line, in terms of cumulative number born, by 28 weeks of age (Wallinga & Bakker, 1978). Under a system of continuous pairing, their high line was unable to sustain its high litter size over successive parities, unless the male was removed to prevent post-partum fertilization.

Few would dispute that an intermediate litter size would be favoured by natural selection, extremes in either direction being less fit. But this statement avoids a more critical question – what determines the level of intermediacy? A glib answer might be that the exact level will depend on the amount of environmental support. But if we are to invoke adaptation on that scale, for a character of low heritability, we need a lot of time in a very constant environment. And even were this so, it would leave the population potentially very vulnerable to environmental change. Extreme adaptation can be self-defeating, and the fittest mouse is probably the one with the largest number of surviving and fertile descendants that could – conceptually – cope with a fairly broad range of environments. Over its evolutionary history, the mouse will have been subjected to natural selection for increased litter size – litters about as large as can be sustained *on average* in a variety of conditions, with some spare capacity in fertility that can be cashed in when conditions are favourable. The genetical evidence is entirely compatible with this – the amount of non-additive variance in the trait reflecting past natural selection for larger litters, but with enough additive variance remaining to show that litter size is not the only component of fitness, and that it could, if need be, still respond to further selection.

CONCLUSIONS

It has been argued that both body weight and litter size, under natural conditions, have been subjected to at least a mild form of stabilizing selection whereby extreme deviants in either direction are

selectively eliminated. But that statement, as it stands, is a trivial one; it is probably true of all traits at all times. A more meaningful question might be: is there a narrow range, or a wide one, over which the effects of natural selection are not easily detectable? The problem here is that we cannot compare different traits on the same scale; there is no logical way of deciding whether corpora lutea counts are more variable than tail length. We can, however, begin to make objective comparisons if we ask how much of the variation is attributable to genetical causes, and to what kind of genetical causes. The answer is clear: body size has a considerable amount of additive variance, giving high estimates for the heritability, whereas fertility traits have lower heritabilities and more non-additive variance. Body weight changes are readily brought about by selection; fertility is more subjected to the effects of inbreeding and crossing. Though most of the evidence comes from laboratory populations, studies on wild mice do not suggest that they are particularly different in their genetic architecture. If we accept the premises set out in the Introduction, then it is at least a reasonable speculation that body weight has not been under the strong influence of natural selection during its evolutionary history. Fertility, on the other hand, has been the subject of considerable natural selection; indeed, if we take fertility and viability together, they *are* natural selection – and there is nothing else that natural selection can be. The component traits of fertility (and litter size is only a component) can still show some additive variance. This is explained, at least in part, by the complex interactions among the components; changes in one component may be buffered by compensating changes in other components, with perhaps little effect on fertility as a global trait. In a sense, the main conclusion was stated at the beginning. What I have attempted in this review is to marshall some of the evidence to support it.

ACKNOWLEDGEMENTS

I am greatly indebted to Dr E. J. Eisen for his patient and friendly discussion of points that arose during the preparation of this manuscript. I am particularly grateful to him for his struggle to read the manuscript in a very rough form – and I apologise to him for dealing selectively with his valuable comments.

I wish to thank Professor R. J. Berry and Dr Carol B. Lynch for drawing my attention to some critical references. To them, and particularly to the authors concerned, I also apologise for any omissions.

REFERENCES

Bakker, H. (1974). Effect of selection for relative growth rate and bodyweight of mice on rate, consumption and efficicency of growth. *Meded. LandbHoogesch. Wageningen* 1974: 74–78.

Bakker, H., Wallinga, J. H. & Politiek, R. D. (1978). Reproduction and body weight of mice after long-term selection for large litter size. *J. Anim. Sci.* 46: 1572–1580.

Barnett, S. A., Munro, K. M. H., Smart, J. L. & Stoddart, R. C. (1975). House mice bred for many generations in two environments. *J. Zool., Lond.* 177: 153–169.

Bateman, N. (1966). Ovulation and post-ovulational losses in strains of mice selected from large and small litters. *Genet. Res.* 8: 229–241.

Batten, C. A. & Berry, R. J. (1967). Prenatal mortality in wild-caught house mice. *J. Anim. Ecol.* 36: 453–463.

Berry, R. J. & Jakobson, M. E. (1975). Adaptation and adaptability in wild-living house mice (*Mus musculus*). *J. Zool., Lond.* 176: 391–402.

Berry, R. J., Jakobson, M. E. & Peters, J. (1978). The house mice of the Faroe Islands: a study in microdifferentiation. *J. Zool., Lond.* 185: 73–92.

Berry, R. J., Peters, J. & Van Aarde, R. J. (1978). Sub-Antarctic house mice: colonization, survival and selection. *J. Zool., Lond.* 184: 127–141.

Bradford, G. E. (1968). Selection for litter size in mice in the presence and absence of gonadotrophin treatment. *Genetics* 58: 283–295.

Bradford, G. E. (1969). Genetic control of ovulation rate and embryo survival in mice. I. Response to selection. *Genetics* 61: 905–921.

Bradford, G. E. (1971). Growth and reproduction in mice selected for rapid body weight gain. *Genetics* 69: 499–512.

Coleman, D.L. (1978). Diabetes and obesity: shifty mutants? *Nutr. Rev.* 36: 129–132.

Connor, J. L. & Bellucci, M. J. (1979). Natural selection resisting inbreeding depression in captive wild house mice (*Mus musculus*). *Evolution* 33: 929–940.

Ebert, P. D. & Hyde, J. (1976). Selection for agnostic behaviour in wild female *Mus musculus*. *Behav. Genet.* 6: 291–304.

Eisen, E. J. (1974). The laboratory mouse as a mammalian model for the genetics of growth. *Proc. World congr. Genet. Appl. Livestock Prod.* I: 467–492.

Eisen, E. J. (1976). Results of growth curve analyses in mice and rats. *J. Anim. Sci.* 42: 1008–1023.

Eisen, E. J. (1978). Single-trait and antagonistic index selection for litter size and body weight in mice. *Genetics* 88: 781–811.

Eisen, E. J., Bakker, H. & Nagai, J. (1977). Body composition and energetic efficiency in two lines of mice selected for rapid growth rate and the F_1 crosses. *Theor. appl. Genet.* 49: 21–34.

Eisen, E. J. & Bandy, T. (1977). Correlated responses in growth and body composition of replicated single-trait and index selected lines of mice. *Theor. appl. Genet.* 49: 133–144.

Eisen, E. J., Hanrahan, J. P. & Legates, J. E. (1973). Effects of population size and selection intensity on correlated responses to selection for postweaning gain in mice. *Genetics* 74: 157–170.

Falconer, D. S. (1955). Patterns of response in selection experiments with mice. *Cold Spring Harbor Symp. Quant. Biol.* No. 20: 178–196.

Falconer, D. S. (1960a). Selection of mice for growth on high and low planes of nutrition. *Genet. Res.* 1: 91–113.

Falconer, D. S. (1960b). The genetics of litter size in mice. Symposium on mammalian genetics and reproduction. *J. cell. comp. Physiol.* 56 (Suppl. 1): 153–167.

Falconer, D. S. (1963). Qualitatively different responses to selection in opposite directions. In *Statistical genetics and plant breeding:* 487–490. Hanson, W. D. & Robinson, H. F. (Eds). (Publ. 982 Nat. Acad. Sci.) Washington D. C.: Nat. Res. Council.

Falconer, D. S. (1971). Improvement of litter size in a strain of mice at a selection limit. *Genet. Res.* 17: 215–235.

Falconer, D. S. (1973). Replicated selection for body weight in mice. *Genet. Res.* 22: 291–321.

Falconer, D. S., Gauld, I. K. & Roberts, R. C. (1978a). Cell numbers and cell sizes in organs of mice selected for large and small body size. *Genet. Res.* 31: 287–301.

Falconer, D. S., Gauld, I. K. & Roberts, R. C. (1978b). Growth control in chimaeras. In *Genetic mosaics and chimaeras in mammals.* Russel, L. B. (Ed.). New York: Plenum Publishing Corporation.

Fisher, R. A. (1930). *The genetical theory of natural selection.* Oxford: Oxford University Press.

Fowler, R. E. (1962). The efficiency of food utilization, digestibility of foodstuffs and energy expenditure of mice selected for large or small body size. *Genet. Res.* 3: 51–68.

Goodale, H. D. (1938). A study of the inheritance of body weight in the albino mouse by selection. *J. Hered.* 29: 101–112.

Hayes, J. F. & McCarthy, J. C. (1976). The effects of selection at different ages for high and low body weight on the pattern of fat-deposition in mice. *Genet. Res.* 27: 389–403.

Hill, W. G. (1971). Design and efficiency of selection experiments for estimating genetic parameters. *Biometrics* 27: 293–311.

Kimura, M. (1957). Some problems of stochastic processes in genetics. *Ann. math. Statist.* 28: 882–901.

Land, R. B. (1970). Genetic and phenotypic relationships between ovulation rate and body weight in the mouse. *Genet. Res.* 15: 171–182.

Land, R. B. & Falconer, D. S. (1969). Genetic studies of ovulation rate in the mouse. *Genet. Res.* 13: 25–46.

Lang, B. J. & Legates, J. E. (1969). Rate, composition and efficiency of growth in mice selected for large and small body weight. *Theor. appl. Genet.* 39: 306–314.

Laurie, E. M. O. (1946). The reproduction of the house mouse (*Mus musculus*) living in different environments. *Proc. R. Soc.* (B) 133: 248–281.

Lerner, I. M. (1954). *Genetic homeostasis.* Edinburgh: Oliver & Boyd.

Lynch, C. B. (1977). Inbreeding effects upon animals derived from a wild population of *Mus musculus. Evolution* 31: 526–537.

MacArthur, J. W. (1944). Genetics of body size and related characters. I. Selecting small and large races of the laboratory mouse. *Am. Nat.* 78: 142–157.

McCarthy, J. C. & Doolittle, D. P. (1977). Effects of selection for independent

changes in two highly correlated body weight traits of mice. *Genet. Res.* 29: 133–145.

McPhee, C. P. & Neill, A. R. (1976). Changes in body composition of mice selected for high and low eight week weight. *Theor. appl. Genet.* 47: 21–26.

Monteiro, L. S. & Falconer, D. S. (1966). Compensatory growth and sexual maturity in mice. *Anim. Prod.* 8: 179–192.

Pidduck, H. G. & Falconer, D. S. (1978). Growth hormone function in strains of mice selected for large and small size. *Genet. Res.* 32: 195–206.

Plomin, R. J. & Manosevitz, M. (1974). Behavioural polytypism in wild *Mus musculus*. *Behav. Genet.* 4: 145–157.

Poiley, S. M. (1972). Growth tables for 66 strains and stocks of laboratory animals. *Lab. Anim. Sci.* 22: 757–779.

Priestley, G. C. & Robertson, M. S. M. (1973). Protein and nucleic acid metabolism in organs from mice selected for larger and smaller body size. *Genet. Res.* 22: 255–278.

Radcliffe, J. D. & Webster, A. J. F. (1976). Regulation of food intake during growth in fatty and lean female Zucker rats given diets of different protein content. *J. Nutr.* 36: 457–469.

Rahnefeld, G. W., Comstock, R. E., Boylan, W. J. & Singh, M. (1965). Genetic correlation between growth rate and feed per unit gain in mice. *J. Anim. Sci.* 24: 1061–1066.

Roberts, R. C. (1961). The lifetime growth and reproduction of selected strains of mice. *Heredity* 16: 369–381.

Roberts R. C. (1965a). Some contributions of the laboratory mouse to animal breeding research. Part I. *Anim. Breeding Abstr.* 33: 339–353.

Roberts, R. C. (1965b). Some contributions of the laboratory mouse to animal breeding research. Part II. *Anim. Breeding Abstr.* 33: 515–526.

Roberts, R. C. (1966). The limits to artificial selection for body weight in the mouse. II. The genetic nature of the limits. *Genet. Res.* 8: 361–375.

Roberts, R. C. (1967a). Some evolutionary implications of behaviour. *Can. J. Genet. Cytol.* 9: 419–435.

Roberts, R. C. (1967b). The limits to artificial selection for body weight in the mouse. III. Selection from crosses between previously selected lines. *Genet. Res.* 9: 73–85.

Roberts, R. C. (1974). Selection limits in the mouse and their relevance to animal breeding. *Proc. World Congr. Genet. appl. Livestock Prod.* 1: 493–509.

Roberts, R. C. (1979). Side-effects of selection for growth in laboratory animals. *Livestock Prod. Sci.* 6: 93–104.

Roberts, R. C., Falconer, D. S., Bowman, P. & Gauld, I. K. (1976). Growth regulation in chimaeras between large and small mice. *Nature, Lond.* 260: 244–245.

Robertson, A. (1960). A theory of limits in artificial selection. *Proc. R. Soc.* (B), 153: 234–249.

Robinson, D. W. & Bradford, G. E. (1969). Cellular response to selection for rapid growth in mice. *Growth* 33: 221–229.

Shire, J. G. M. (1976). The forms, uses and significance of genetic variation in endocrine systems. *Biol. Rev.* 51: 105–141.

Sutherland, T. M., Biondini, P. E., Haverland, L. H., Pettus, D. & Owen, W. B. (1970). Selection for rate of gain, appetite and efficiency of feed utilization in mice. *J. Anim. Sci.* 31: 1049–1057.

Sutherland, T. M., Biondini, P. E. & Ward, G. M. (1974). Selection for growth rate, feed efficiency and body composition in mice. *Genetics* 78: 525–540.

Timon, V. M. & Eisen, E. J. (1970). Comparisons of *ad libitum* and restricted feeding of mice selected and unselected for postweaning gain. I. Growth, feed consumption and feed efficiency. *Genetics* 64: 41–57.

Timon, V. M., Eisen, E. J. & Leatherwood, J. M. (1970). Comparisons of *ad libitum* and restricted feeding of mice selected and unselected for postweaning gain. II. Carcase composition and energetic efficiency. *Genetics* 65: 145–155.

Wallinga, J. H. & Bakker, H. (1978). Effect of long-term selection for litter size in mice on lifetime reproduction rate. *J. Anim. Sci.* 46: 1563–1571.

Yanai, J. & McClearn, G. E. (1972a). Assortative mating in mice and the incest taboo. *Nature, Lond.* 238: 281–282.

Yanai, J. & McClearn, G. E. (1972b). Assortative mating in mice. I. Female mating preference. *Behav. Genet.* 2: 173–183.

PAPER 29.

Lessons to be drawn from selection experiments with laboratory animals.

Proc. World Congress on Sheep and Beef Cattle Breeding, Palmerston North and Christchurch, New Zealand, 1980. Vol. 1, 253-259.
Eds. R.A. Barton and W.C. Smith. Palmerston North, NZ: Dunmore Press. 1982

by

R.C. ROBERTS.

From: PROCEEDINGS OF THE WORLD CONGRESS ON SHEEP AND BEEF CATTLE BREEDING
Volume I: Technical

First published 1982
Edited by: R.A. Barton and W.C. Smith

Published by The Dunmore Press Limited, P.O. Box 5115, Palmerston North, New Zealand

26 LESSONS TO BE DRAWN FROM SELECTION EXPERIMENTS WITH LABORATORY ANIMALS

R.C. Roberts
ARC Unit of Animal Genetics, Institute of Animal Genetics, West Mains Road, Edinburgh EH9 3JN, Scotland

SUMMARY

Responses to selection are readily obtained for a variety of traits related to fertility and growth, though a combination of the two-total biomass at weaning time – appears to be singularly recalcitrant. Laboratory populations are often of small size, leading to some inbreeding. Experience suggests that inbreeding should not exceed 1.5% per generation, if a decline in fertility is to be avoided. Correlated responses are invariably observed in traits other than the one selected. If a research programme aims to estimate such correlated responses, the need to replicate the selection is stressed.

INTRODUCTION

The advantages of small animals in breeding research in terms of economy of time and cost are widely appreciated, and reviews by Roberts (1965a, b) and by Eisen (1974) illustrate their contributions. The relevance of laboratory animals to animal science generally is well established, but for animal breeding in particular, small animals (even *Drosophila*) have an especially important role in the training of future animal breeders. Only with laboratory animals can a student actually conduct a breeding programme over several generations, and there is no substitute for the experience that such a programme provides: the logistics of getting the right animals in the right numbers to the right place at the right time, to make real decisions on what to mate to what; to take emergency action when fertility or viability problems beset a well-laid scheme and, most of all, to sample the rewards and disappointments of honest endeavour.

THE LIMITATIONS OF LABORATORY EXPERIMENTS

Laboratory animals, by their nature, are of limited interest, and the world is not going to benefit significantly if more mice are produced or are made to grow

faster. The results, to be useful, have to carry over to domestic livestock by generating higher yields, or lower costs. There are obvious difficulties about translating results across species; the mouse is not a ruminant, it does not need to forage, nor does it need to meet the idiosyncrasies of the consumer market. But more than that, there are peculiarities of laboratory animal experimentation that may have little application to the management of domestic livestock.

Laboratory animals are normally kept in very uniform environments, usually of constant ambient temperature and often of regular light. They receive a high protein diet, of constant chemical composition, usually fed *ad libitum*, and the same diet is used for breeding and growing stock. The mating structure is unlike that of farm animals; pair matings are widely used, or at most, only a very few dams mated to each sire. Male castrates hardly ever figure in the evaluation of the results. Any of these may prejudice direct application to commercial conditions. But the biggest difference lies in the objectives of a laboratory breeding programme – usually a single trait, or at best a combination of 2 traits. No breeder of commercial livestock, aiming for high yields at low cost of a marketable product, can afford to be that single-minded.

POTENTIAL CONTRIBUTIONS OF LABORATORY EXPERIMENTS

With this in mind, where do laboratory animals fit into animal breeding research? Their value in training future animal breeders has already been noted, and this is a continuing need. The second area of dependency on laboratory species is that of the validation of quantitative genetic and animal breeding theory. However, this contribution is now largely historical. There is enough validated theory, well enough understood, to serve the needs of animal improvement programmes for some time to come; for that, mice and fruit flies can be thanked.

The main contribution of laboratory species, as applied to animal breeding research, will come in the future from the more detailed analyses of complex phenotypes. Take growth as the prime example; readers interested in fertility, or lactation, or wool can do their own adaptation. The complexity of growth as a phenotype is enormous. How can genetic variation in appetite be explained, for a start? What are the physiological and biochemical variables at this level, and how are they affected by genes? Then, what does the animal do with the ingested food? Is there genetic variation in digestion? As a matter of fact, there may not be, but if not, why not? Next comes the big one – what does the animal do with its pool of digested metabolites? What does it need to maintain itself, with respect to body temperature and various metabolic functions, and to continuously replace its tissues? How can genetic variation be partitioned at this level, and what is genetically correlated with what? Then, what does the animal do with the remainder, and how does it assign its priorities? For that matter, how does the animal decide on the distinction between maintenance and further growth? Or does it distinguish? What genetic variation is there in the partitioning of metabolites to various destinations?

In short, there are some lids to be taken off some black boxes. If a Jersey and a Hereford cow were given an area of grass, they would do different things with it. Somewhere there is genetic variation, but little else is known. But genetic experiments with cattle are long-term and expensive. Mouse experiments are short-term and inexpensive. Much basic physiology is common to all mammals, and biochemistry is standard. So there is some idea of where it may be possible to extrapolate and where it may not. Therefore laboratory animals can be used to characterize the basic genetical properties of complex systems and it ought to be possible to use such genetic knowledge to optimize genetic manipulations. It should be easier to rationalize various breeding schemes if it were known what sources of genetic variation are at the disposal of the breeder.

This discussion stops short of advocating that animal breeders cash in on the techniques of the molecular biologist and get a complete profile of every gene action in every metabolic pathway. Not because the idea is too futuristic, nor because the results would lack interest, but because the information would be of little use in a practical breeding programme. The question is what *effects* genes have, and where. What the genes actually do, and how, is something else – and that knowledge would not help the research worker to reshuffle them.

RESPONSES TO SELECTION

The first and obvious lesson, amply illustrated in the reviews cited earlier, is that selection usually yields a response in the desired direction for a variety of traits related to growth and fertility. However, this cannot be assumed as the inevitable outcome, and one of the failures to get a response has serious implications for farm animals. This is the total weight of the litter at weaning time. No response was observed on either of 2 diets by Dalton and Bywater (1963), and the results of Roberts and Steane (unpubl.) are also negative. Similarly, Bateman (reported by Falconer, 1955) failed to increase 12-day litter weight, which is about peak lactation for the mouse. Increases in 12-day litter weight were reported by Eisen *et al.* (1970), Robinson *et al.* (1974) and by Nagai *et al.* (1978), but in all of these studies litter size had been standardized, which in effect excludes one of the variables that the breeder might wish to improve. Total litter weight is a trait of inordinate complexity, comprising genetic and maternal effects on the 2 components – number and individual weight – with a multitude of covariances among them. The issues are examined by Eisen (1981) who developed indices to optimize selection responses under various procedures. There is urgent need for experimentation to explore this topic further.

Having noted the failure, there is a voluminous literature illustrating success, and it should be noted that responses can readily be obtained with small populations, often with effective population sizes of no more than 30 or 40. The effective number is, of course, related to the harmonic mean of the numbers in the two sexes, and an effective population size of 32, for instance, is unattain-

able with fewer than 8 sires, irrespective of the number of dams. But a modest sheep flock of 10 rams and 500 ewes would give an effective population size of about 40, and laboratory animal experience suggests that it is possible to select successfully in a population of that size.

To the extent that the breeding objective could be limited to a single trait, with a reasonable heritability, experience further suggests that the programme would not reach a limit for 10 to 20 generations. Over this length of time it could be suggested as an approximate guide, that traits like litter size could possibly be improved by 30% and body weight, or rate of gain, by as much as 50%, so there are worthwhile gains to be made. But the total selection would have to be concentrated on the one trait and not dissipated among others. If several traits are to be combined in one index, the premises of the argument are not affected. The difference would be that on an index, the breeder would perhaps not notice startling changes in component traits, while his overall phenotype (measured by his monetary return) would be hard to assess for extraneous economic reasons.

There are 2 reasons to be wary of small populations. The first, is the accumulation of inbreeding with the consequent decline in litter size, at the rate of about ½ mouse per 10% increase in inbreeding under continued full-sib mating. However, if inbreeding proceeds slowly, natural selection may counteract at least some of it by (presumably) weeding out the more homozygous, and thus less fertile, individuals. Falconer (1960) reported that an unselected control strain, of effective size of 40, leading to inbreeding at the rate of 1.25% per generation, accumulated an inbreeding coefficient of 32% without any decline in litter size. In contrast, the average of 6 control strains in the same laboratory, each strain of effective size of 32 and with a rate of inbreeding of 1.56% per generation, showed a decline of 1½ mice per litter by the time the inbreeding coefficient had reached 30% (Falconer, 1973). This is precisely the rate predicted from more rapid inbreeding. The conclusion from laboratory mice is therefore clear: effective population sizes in the range of 30 to 40 are marginal, if the cumulative effects of inbreeding are to be avoided. As an empirical rule, breeding schemes where the rate of inbreeding exceeds 1.5% per generation might be at some risk.

The second reason for using larger populations is equally widely appreciated, namely, that the total genetic gain is likely to be greater. This effect has been well illustrated with *Drosophila*, both in the short-term (Frankham *et al.*, 1968) and in the long-term (Jones *et al.*, 1968), while Hanrahan *et al.* (1973) found the same trend with mice. Thus, the experimental evidence supports the theoretical advantages of increasing the population size.

The conclusion from this section is that while individual breeders with limited facilities could certainly aspire to some genetic gains within their own flocks or herds, they are unlikely to compete effectively with those who operate on a larger scale. Dairy cattle breeding is already largely based on national or regional schemes, while large international companies have taken over poultry breeding, with pigs going the same way. Thus far, the breeding of sheep and of beef cattle has been less centralized. But there are sound genetic reasons support-

ing the trend towards co-operative ventures where resources are pooled and which at the same time allow individual breeders to retain something of their identity and their active participation in the improvement programme.

THE RELIABILITY OF SELECTION RESPONSES

A particular result from a selection experiment – like any other experiment – is a demonstration of one possible outcome. If it were to be repeated on the same numbers selected in the same way under the same conditions, the outcome could be different. This is well illustrated by Falconer (1973), who selected from the same base population 6 lines of mice for high body weight and 6 for low body weight. By generation 10, the 6 high lines varied in mean from 26 to 32 g and the 6 low ones from 16 to 19 g. Each time parents are selected, the gene pool is sampled, and accidents of sampling (genetic drift) lead to cumulative differences. Hill (1974) examined the consequences for a selection programme.

These considerations have one serious repercussion in the practical situation where selected lines are being compared, or selection methods evaluated. It is still common practice to compare 2 selected lines, using the within line variance as the source of error – an approach branded by Hill (1980) as 'naive'. Allowance should be made for the fact that the 2 lines may well differ for reasons of drift alone, and it is wrong to use an inappropriate error term to draw inferences about the populations from which the lines were sampled, or the selection methods which generated them. The direct method to assess the drift variance is by replicating the selection, though Hill (1980) has given an approximate formula for estimating it in the absence of replication. This formula applies, however, only to the trait under selection, and Hill notes further that if the objective is to estimate correlated responses, replication is necessary before the sampling variances of those correlated responses can be measured.

The conclusion from this section is that whereas replication may perhaps be of little interest in a practical improvement scheme, it should be a central feature of animal breeding research programmes, if any valid inferences are to be drawn.

CORRELATED RESPONSES TO SELECTION

This topic has been reviewed elsewhere (Roberts, 1979) and there is room here only for the main conclusions. It is well known that when one trait is selected, others change in consequence. As a particular example, selection for increased weight, or rate of gain, in laboratory mice almost always has two undesirable consequences. First, the large mice get fat, especially at older ages, as illustrated most recently by Allen and McCarthy (1980). Second, their fertility declines to the point where they become difficult, if not impossible, to maintain (see Roberts, 1967a, for a specific instance). The implications are that breeders seeking more rapid rates of gain must be prepared to manage the breeding animals (females especially) in such a way that they do not accumulate excess

fat. It is true that undesirable traits might be incorporated, with a negative weighting, in an index, but where possible, managemental control is to be preferred. As pointed out by Eisen and Bandy (1977), antagonistic index selection carries a cost in terms of reduced genetic gains.

SELECTION LIMITS

Selection responses do not continue indefinitely, if only because ultimately the genetic variance in the population becomes exhausted. However, this is not an immediate problem in animal breeding, and the writer has nothing new to add since an earlier review (Roberts, 1974). However, experience with laboratory animals has one lesson. If an improved strain, or breed is to be outcrossed (a hypothetical example might be to seek resistance to some disease), then it is unacceptable to cross it to genetically-inferior material. It takes too long to recover the initial level. More than that, if a new population is to be constructed by crossing existing breeds or strains, the higher the initial level of performance, the greater the subsequent response (Roberts 1967a, b). It is not worth sinking effort into trying to upgrade inferior material.

GENERAL CONCLUSION

Robertson (1958) claimed that, as a result of his work with *Drosophila,* he felt far more competent to discuss improvement of dairy cattle than if all his time had been spent analysing milk records and perhaps breeding a few cattle. Some workers may display less alacrity in crossing species boundaries, but his general point can be paraphrased thus: work with laboratory animals does not constrain the scientist to lift the mean performance, but provides a more general context in which to explore problems and ideas. It is from a deeper understanding of the issues, rather than from any instant solutions, that laboratory animal work has been of greatest value.

REFERENCES

Allen, P. and McCarthy, J.C., 1980. Anim. Prod., *31:* 1.
Dalton, D.C. and Bywater, T.L., 1963. Anim. Prod., *5:* 317.
Eisen, E.J., 1974. Proc. 1st World Congr. Genetics Applied to Livest. Prod., Vol. I, Madrid, Spain. p. 467.
Eisen, E.J., 1981. Z. Tierzüchtg Züchtungsbiol., *98:* 55.
Eisen, E.J. and Bandy, T., 1977. Theoret. Appl. Genet., *49:* 133.
Eisen, E.J., Legates, J.E. and Robinson, O.W., 1970. Genetics, *64:* 511.
Falconer, D.S., 1955. Cold Spring Harb. Symp. Quant. Biol., *20:* 178.
Falconer, D.S., 1960. J. Cell. Comp. Physiol., *56* (Suppl. 1): 153.
Falconer, D.S., 1973. Genet. Res., Camb., *22:* 291.
Frankham, R., Jones, L.P. and Barker, J.S.F., 1968. Genet. Res., Camb., *12:* 237.

Hanrahan, J.P., Eisen, E.J. and Legates, J.E., 1973. Genetics, *73:* 513.
Hill, W.G., 1974. Biometrics, *30:* 363.
Hill, W.G., 1980. In: *Selection Experiments in Laboratory and Domestic Animals.* Ed. A. Robertson. Proc. Symp. Harrogate, England, 1979. Pub. Commonwealth Agric. Bureau, Slough, England. 245 pp.
Jones, L.P., Frankham, R. and Barker, J.S.F., 1968. Genet. Res., Camb., *12:* 249.
Nagai, J., Eisen, E.J., Emsley, J.A.B. and McAllister, A.J., 1978. Genetics, *88:* 761.
Roberts, R.C., 1965a. Anim. Breed. Abst., *33:* 339.
Roberts, R.C., 1965b. Anim. Breed. Abst., *33:* 515.
Roberts, R.C., 1967a. Genet. Res., Camb., *9:* 73.
Roberts, R.C., 1967b. Genet. Res., Camb., *9:* 87.
Roberts, R.C., 1974. Proc. 1st World Congr. Genetics Applied to Livest. Prod., Vol. I, Madrid, Spain. p. 493.
Roberts, R.C., 1979. Livest. Prod. Sci., *6:* 93.
Robertson, A., 1958. Proc. Xth Int. Congr. Genet., Vol. I, Montreal, Canada. p. 199.
Robinson, W.A., White, J.M. and Vinson, W.E., 1974. Theoret. Appl. Genet., *44:* 337.

PAPER 30

Genes with large effects - theoretical aspects in livestock breeding.

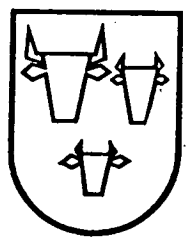
Proc. 2nd World Cong. appl. to Anim. Prod., Vol. 6, 420-438.
Madrid. 1982.

by

R.C. ROBERTS & C. SMITH.

2nd. WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION

II CONGRESO MUNDIAL DE GENETICA APLICADA A LA PRODUCCION GANADERA



6 ROUND TABLES
MESAS REDONDAS

4th-8th OCTOBER 1982
4-8 OCTUBRE 1982

MADRID
SPAIN - ESPAÑA

GENES WITH LARGE EFFECTS - THEORETICAL ASPECTS IN LIVESTOCK BREEDING

Genes con efectos básicos: aspectos teóricos en mejora animal

R.C. ROBERTS*
C. SMITH*

SCOTLAND (U.K.)

Introduction

Animal breeding theory leans heavily on the concepts and methodology of quantitative genetics, which themselves were developed on the assumption of many genes with small effects. At least until recently, there was no cause to question the validity of this assumption with respect to production traits in domestic livestock. The possible identification of major genes influencing production was nevertheless kept under continual review, and received some impetus from studies of polymorphisms in general and of blood antigens in particular. However, there has been little room to challenge the generality of a conclusion by Neimann-Sørenson and Robertson (1961), in their case with specific reference to blood groups and dairy production that such approaches had a low predictive value i.e. the influence of specific loci on production could not readily be detected. However, this does not mean that genes with major effects do not exist, and recent work on the halothane and K-88 loci in pigs and on the Booroola gene in sheep suggests the need to re-examine the possibilities.

Historically, single genes have found two applications in animal breeding, neither of which was directly connected with production. The first was probably adventitious, but in the event, different alleles at a few loci became the trade marks of various breeds, chiefly colour variants and horns. The second application was in the control of congenital abnormalities frequently reviewed in the past (see, for instance, Young, 1967 and Lauvergne, 1968). These are not discussed further here.

The only class of livestock where single genes have had any impact on the commercial product is poultry. The use of sex-linked genes to separate chicks by sex is a well-known text-book classic. A recessive dwarf mutant that can reduce maintenance costs in female broilers has been examined by Mérat

* ARC Animal Breeding Research Organisation. West Mains Road. Edinburgh EH9 3JQ. Scotland. U.K.

and Ricard (1974) and is now in commercial use. Several other mutants in chickens, particularly those concerning skin colour, egg colour and feathering rates, have been exploited to commercial advantage. Somes (1978) provides a comprehensive listing of such genes in poultry.

Despite the fact that genes with large effects on production traits do not feature prominently in cattle, sheep and pigs, the evidence from laboratory animals suggests that they may yet be found. In the laboratory mouse, for instance, obese, adipose and diabetes are all mutants increasing body weight, while dwarf and pygmy reduct it. In all these cases the effect on weight is sufficient to identify homozygotes unambiguously. It may be no accident that several of these mutants were found in lines selected for weight in the direction of the mutant effect. Mutants affecting the selected phenotype, and which are lethal when homozygous, have been found almost routinely in Drosophila since they were first reported by Clayton, Morris and Robertson (1957). As domestic livestock become more intensely selected, the expectation must be that mutants affecting the selected traits may be found. This implies some effect in the heterozygote, even among genes commonly classified as "recessive". The laboratory evidence also indicates a potential source of trouble: the Drosphila mutants tend to be lethal, and the mouse mutants sterile, in the homozygous state. In either event, the discovery of similar mutants in livestock would reduce their appeal, even if some of their effects were beneficial.

The appearance of such mutants in selected lines in laboratory animals serves to remind us that any major gene can be exploited by the standard methods of quantitative genetics. It is true that they violate some of the assumptions, but the methods are generally robust enough to withstand this. At one level, it might therefore be a defensible argument

that we need not trouble about genes of large effect; if they are there, and if they are beneficial, they will be selected and moved to the desired frequency, or fixed. However, if segregants at a single locus can be identified, and particularly if there is dominance at that locus, quantitative methods are less efficient in the exploitation of such genes than would be the mating of known genotypes. Any overdominance would render selection inappropriate to capitalize on the advantages.

Against this background, we shall examine the problems of identifying genes with large effects, their effect on genetic parameters, their optimal use in breeding programmes, and the information required for their effective utilization.

Identifying genes with large effects

As already mentioned, selection will operate on all segregating loci, including those with large effects. To that extent, the existence of major genes is primarily of heuristic interest. However, there are some practical implications in special cases, mostly those involving overdominance. In this case selection will lead to an equilibrium gene frequency, maintained by continuing selection pressure but without yielding further genetic response at the locus. This could arise from overdominant effects of the locus for a major trait, or if there were antagonistic effects of the gene at different stages of the life cycle (e.g. favourable when animals are selected and unfavourable when the selected animals reproduce). Alternatively, the pleiotropic effects of the gene may be beneficial for some traits and harmful for others, the selection objective being a compound of both kinds of traits. While these practical considerations are not limited to major genes, their consequences are more serious with genes of large effect. Knowledge about

the presence of the gene and its diverse effects would allow different breeding strategies to exploit its benefits and minimise its drawbacks. Thus there might be a case for fixing the different alleles in different lines and to use cross-breeding to generate the commercial animals. Similarly, the development of sire and dam lines might be indicated, to maximise reproductive performance while at the same time deriving some benefit for growth and carcass traits.

It is necessary to define what we mean by a major gene, although in a continuity of gene effects, the definition is quite arbitrary. Following Morton and Maclean (1974) we may think of a major locus as one having an effect of at least one standard deviation of the metric trait, as measured by the difference between the two homozygotes ($2a \geq \sigma$, as defined later). With genes of very large effect ($2a \geq 2\sigma$) segregation and kurtosis may be apparent in the population, but as the effect falls, segregation and kurtosis will become harder to detect (Piper, 1972). Thus on our definition of a major locus, many such loci will not be suspected in the population.

It would be of some practical utility to know the distribution of gene effects in the population, and to know what proportion of the genetic variation is due to 'major genes'. Much of the work on the number of loci controlling a metric trait is attributable to Prof. A. Robertson and his students at Edinburgh, particularly Piper (1972) and Shrimpton (1981), extending the methods of Breese and Mather (1957) and Thoday (1961) for locating genes on the chromosomes of Drosophila, with respect to bristle number. Cumulatively, these studies have shown a large number of loci to be involved and with a range of gene effects, some of them quite large effects. However, it becomes virtually impossible to define a locus by these techniques, and the effects would

be more accurately assigned to chromosomal segements, which may themselves represent a group of closely linked genes. The distinction between loci and chromosomal segments is even more intractable with mammalian species. Not only do such species lack the detailed linkage map available in Drosphila, but they lack also the techniques for suppressing crossing-over to render possible the examination of individual chromosomes. For domestic livestock, or even laboratory mammals, we are not within sight of any method of distinguishing single genes from a cluster of linked genes. That includes the fragmentation of DNA strands with restriction enzymes, though the technology is now sufficiently advanced to make this potentially feasible in a well-mapped species. However, it may be of more academic than of practical interest to distinguish a single locus from a small chromosomal segment, because the consequences in both cases are similar.

Single genes have been identified and utilised in plant breeding work, particularly in the areas of disease resistance, biochemical variants, (e.g. high lysine corn) and agronomic traits (e.g. dwarf type cereals). These genes have been discovered by screening large populations and frequently among a diversity of genetic material. Such searches do not seem to have been conducted in animals, or if they have, seem to have been unsuccessful. In practice, extensive screening of animal populations pose problems both in design and execution, since the approach would be to identify extreme variants and then to establish the genetic features of those variants. Even if proved to be a single gene, its incorporation into an improved commercial strain might well lead to an intolerable loss of production, because the effects of the remainder of the improved strain's genome would be diluted by crossing. From this point of view, we may have to await the perfection of techniques for

incorporating small fragments of "foreign" DNA into the genome of an improved strain. That would retain the remainder of the genome intact. While this approach at the moment is probably a little futuristic, the technical advances in that area are sufficiently rapid that they cannot be dismissed as fantasies.

If genetic engineering is to become a practical proposition in domestic livestock, there is a prior need to concentrate on locating specific loci, or specific chromosomal regions, that affect production traits. With this objective in mind, there are two complementary approaches. The first is to screen widely for new variatns, as just mentioned. The second is to locate major gene effects among genomes currently used for production, linkage relationships being exploited as appropriate. Several methods have been proposed in the literature for detecting genes with large effects. These include:

(i) Comparing the distribution of parental lines with those of F_1, F_2 and backcross generations (e.g. Stewart 1969).

(ii) Repeated backcrossing with selection, followed by selfing — applicable only in plants (e.g. Wehrhahn and Allard 1965).

(iii) Detection through linkage with marker loci (Jayakar, 1970, Haseman and Elston, 1972, Geldermann, 1975, Hill, 1975).

(iv) Path and segregation analysis for quantitative traits (Morton and MacLean, 1974, Rao, Morton, Lalouel and Lew, 1979).

(v) Extended family pedigree analysis (Go, Elston and Kaplan, 1978).

(vi) Distribution of differences from mid-parental values (Karlin, Carnelli and Williams, 1979).

While each of these methods may find particular applications, their statistical power depends greatly on the magnitude of the effect of the gene and on its frequency in the population. In general, the power declines

rapidly as the magnitude of the effect falls (see, for example, Robertson, 1973). The methods may therefore be of more value in confirming suspicions of a major gene from other evidence, rather than in establishing the presence of such a gene where none had been expected.

For Mendelian loci, Davie (1979) has developed a simplified method of segregation analysis which may find considerable application in animal breeding. Models for detecting two major loci affecting a quantitative trait were considered by Merry, Roger and Curnow (1979). The resolution of such models, particularly if penetrance is incomplete and variable, may demand maximum likelihood methods (e.g. Smith and Bampton 1976).

The conclusions from this section is that despite many methods and proposals, the detection of major genes in domestic livestock still presents methodological problems. Though adventitious and haphazard in nature, we should perhaps not ignore possible pleitrophic effects on production traits of genes discovered by other routes. A prime example is the gene for halothane reaction in pigs (Eikelenboom and Minkema, 1974) with its well-known association with porcine stress syndorme and its effects on fertility and lean content. Thus far, this somewhat primitive approach to single gene effects has proved to be the most rewarding, but its limitations are self-evident.

Effects on genetic parameters

The effect of a major gene on genetic parameters is most easily visualised as follows. Consider first the genetic parameters in a population in the absence of the major gene, and then introduce into the population

an allele at one locus with a large effect on the trait under consideration. For simplicity, assume that all the variance due to the other loci is unaffected by the introduction of the major gene, and that the environmental variance also remains unaffected.

Following Falconer's (1981) terminology, we may symbolise the properties of the major locus as follows:

Genotype	A_1A_1	A_1A_2	A_2A
Frequency	p^2	$2pq$	q^2
Genotypic value	a	d	-a

If we symbolise the additive genetic and the phenotypic variances as V_A and V_p, respectively, then the heritability(h^2) in the presence of the major gene becomes

$$h_2 = \frac{V_A + 2pq\alpha^2}{V_p + 2pq\alpha^2 + (2\ pqd)^2}$$

where $\alpha = a + d\ (q-p)$.

The effect on the heritability, compared to the usual V_A/V_P in the absence of the gene, is therefore a function of q, a and d. The influence of various values of a and d, over the full range of gene frequency, is shown by Smith and Webb (1981), with the following general conclusions:

(i) The effect of a major locus generally is to increase the heritability; the only exception is in the presence of overdominance at the locus, and then only for a limited range of gene frequency.

(ii) In the absence of dominance, the increase is symmetrical about the value of $q = \frac{1}{2}$, and is proportional to the square of the gene effect (a^2).

(iii) In the presence of dominance, the increase is asymmetrical, the maximum being displaced towards lower frequencies of the dominant allele.

Consonant with the changes in heritability, the response to selection will also be affected in the presence of a major gene. For a given selection intensity, the response is proportional to $V_A/\sqrt{V_P}$. Taking the square-root of the denominator therefore magnifies the effect in the expression given above. The exact gene frequency at which the maximum response is obtained will depend on the dominance or recessiveness of the desired allele. The general conclusion is that the presence of a major gene would enhance the response to selection for traits affected by that gene, the enhancement generally being greatest at intermediate frequencies of the major gene.

As is well known for all single gene effects, selection for a recessive leads to ready fixation, while selection against a recessive becomes progressively less efficient as its frequency diminishes. The effect of a single gene in a selection programme therefore depends on its dominance, and whether the object is its fixation or elimination. Smith and Webb (1981) discuss the implications in more detail, with particular emphasis on removing an undesirable recessive from the population. They show clearly the advantages of test mating with homozygous recessive mates, and the desirability of test mating both sexes. To maximise the information, they propose that the various sources of information be formalised through the use of Bayes' Theorem, combining information from pedigrees, marker genes, phenotype and test mating, and including the variables of gene frequency and penetrance. The use of Bayes' Theorem in this context is perhaps less contentious than in most cases, because the prior probabilities are known from

the Mendelian rules. However, any problems with penetrance or viability may effect those prior probabilities and therefore detract from the rigour of the approach.

If a major gene affects two traits, the genetic correlation (r_A) between those traits is augmented by the major gene in a manner analogous to the effect on the heritability. The formula becomes:

$$r_A = \frac{cov_{A_1A_2} + 2pq\alpha_1\alpha_2}{\sqrt{(V_{A_1} + 2pq\alpha_1)\ (V_{A_2} + 2pq\alpha_2)}}$$

where the numerical subscripts refer to the two traits, and cov_{AA} is the additive covariance between them in the absence of the major gene. Even if this covariance were zero in the absence of the gene, the formula shows how the gene's effects could be quantified in terms of the genetic correlation.

The optimum use of a major gene in a breeding programme

The optimum use of a major gene in breeding programmes has been discussed by Nieman-Sørenson and Robertson (1961) and by Smith (1967), and with specific reference to the halothane test in pig improvement, by Webb and Jordan (1979). The first requirement is that the genotypes at the major locus should be identified, because the desired genotypes can then be incorporated into the selection programme to increase the rate of response. The efficiency of this form of direct selection on the gene depends on the amount of the additive genetic variance due to the major locus, expressed as a proportion (R) of the total additive variance (including that due to the locus) in the trait. If we symbolise the total additive genetic variance as $V_{A(T)}$ (to distinguish it from the V_A used earlier), the ratio becomes:

$$R = (2pq\alpha^2 - \frac{V_p}{N})/V_{A(T)}$$

The term V_p/N removes the biases due to sampling errors in estimating the

parameters of the major locus i.e, a, d and q, and where N is the number of animals tested. Where normal selection is effective in the absence of any knowledge of the major gene, the additional information contributes little to the rate of improvement. But the rate can be increased substantially if the heritability is low, or where indirect selection has to be practised e.g. in sex-limited traits. As before, the gain will be maximised at intermediate fequencies, and diminishes rapidly as fixation is approached. The increased rate of response also depends greatly on the accuracy of the estimates of gene effects. Smith (1967) notes that errors of estimation may not only reduce the expected gains but even make them negative.

The main implications of a major gene in terms of breeding strategy occur when the heterozygote may be the favoured genotype. This does not necessarily depend on overdominance, which might exceptionally occur. Much more commonly, the gene might have antagonistic affects on the overall selection objective, beneficial for some traits and deleterious on others. The best combinations of traits could then well be found inthe heterozygote. The optimum strategy might then call for sire and dam lines to be fixed for different alleles, the commercial product being the heterozygote. While this may be feasible for species with a high reproductive rate, like poultry and pigs, the scope may be a more limited one in cattle and sheep. We should then have to think in terms of fixing a breed e.g. a terminal sire breed, for an allele that might well make the propagation of that pure breed more difficult or costly. An alternative would be to develop screening techniques for a segregating population so that, with artificial insemination, appropriate semen for a specified female might be available on demand.

The information required for the effective utilization of single genes

If a gene of large effect is to be used effectively in a breeding programme, it follows from the last section that we need reliable information on the effects of the different genotypes (at the major locus) on all traits of economic importance. We emphasise all, and if nothing else in this review, we wish to draw attention to the magnitude and importance of this task. It seems to us that in the past, the recommendations about the use of a major gene in a breeding programme have ignored this aspect. It has often been assumed, maybe for the lack of critical evidence, that the pleitropic effects of the gene on other commercial traits may be neutral, and that no deleterious effects may occur. If this is not so, the whole selection programme could be seriously misdirected. The problem was put in shart relief by Webb and Jordan (1979), who pointed out that the benefit from improved carcasses in halothane-positive pigs could be outweighed by its negative effects on viability and fertility.

The first requirement is to obtain an estimate of the performance of each genotype for all traits that could affect economic merit. These might be growth rate, food conversion ratio, food intake, killing out percentage, percentage of lean in the carcass, mortality, fertility, as well as various structural traits like leg weaknesses, dentition etc. Given this, we could estimate the economic value (A) of genotype k as:

$$A_k = \sum_i a_i (x_{ik} - \bar{x}_i)$$

where a_i is the economic value of a unit change in trait i, and x_{ik} is the performance of genotype k (at the major locus) for trait i, with a mean over all genotypes of $\bar{x}_i$. The expression could also be adapted to allow for traits with intermediate optima and possibly non-linear economic weights.

These economic values, being compounded over several traits, are obviously subjected to sampling errors, and for reasons stated earlier, it becomes of critical importance to estimate these errors. Possibly the simplest approach is to calculate the empirical standard error, treating each calculated value of A as an observation in its own right. Alternatively the standard statistical procedures of combining information might be applied for calculating the standard error, and for this, it may be adequate to derive the phenotypic correlation matrix relating all traits within genotypes. It should be noted, however, that the correlation matrix itself will have sampling errors, which need to be taken into account in any general assessment of the approach.

Some of the information will be estimated more reliably than the remainder. For instance, production traits like growth and carcass measurements can usually be estimated relatively accurately, because of their higher heritabilities and low coefficients of variation, and more especially because large numbers of individuals will usually be available. Reproductive traits, on the other hand, will usually be assessed on fewer individuals, because the information is limited to breeding animals. In addition, the heritabilities are lower and the coefficients of variation are high. Nor do the complications end there. Sometimes the traits are of an all-or-none kind, scored 0 or 1; this will be true of traits like sterility and mortality. In addition, it may be difficult to attach accurate economic weights to some traits. While everyone would agree that bad legs or poor teeth are undesirable, their exaxt monetary penalty is uncertain, and furthermore may vary among genotypes at the major locus, or among breeds kept under different conditions or systems of management. Should that be the case, the safest estimates may be those derived for the commercial product kept under commercial conditions,

but relevant information would also be of value to breeders of parent stock.

It is therefore clear that before any firm recommendation about the use of a major gene in a breeding programme can be made with any confidence, a substantial amount of prior information is necessary. The collection of that information will inevitably prove costly and time-consuming, and will demand considerable care in its interpretation. It would be idle to pretend that there is any easy alternative to this exercise, and it is the basis on which the successful exploitation of a major gene will inevitably rest.

Discussion

We defined a major gene as one having an effect of one standard deviation or so on a metric trait, but also pointed out the arbitrariness of this definition. There are often problems of identification of such a gene, more especially if the genetic background modifies its expression. We should also note that single genes may be subjected to environmental variation, often ignored in this context though routinely incorporated in any formal treatment of metric traits. But the well-known example of the himalayan rabbit illustrates the potential difficulties. In the cold, the extremities of the himalayan rabbit become black, but at higher temperatures, the rabbit becomes pure white. There may therefore be certain environments where the presence of a major gene could be wholy undetectable. More generally, it can by no means be assumed that the effects of a gene under certain conditions may unconditionally be expected under others. This is nothing more than a special case of genotype x environment interaction.

Environmental effects can easily be confused with background genetic effects. New mutants frequently, perhaps typically, do not segregate cleanly

when they first arise, and their penetrance could be affected not so much by the environment as by the influence of genes at other loci. This was very clearly demonstrated by Fisher (1950) and Bodmer (1960). Both were working on the same polydactylous mutant in the mouse. The recessive homozygote was not unconditionally polydactylous. Fisher isolated alleles at three other loci which together gave a normal phenotype even when the major mutant gene was homozygous. Bodmer, on the other hand, found modifiers at two yet other loci which enhanced the effect on the phenotype, so much so as to give some polydactyly even in the absence of the major gene. In effect the polydactylous gene revealed the effects of its segregation only in an appropriate genetic background. The modifying strength of the background has been apprectiated since Waddington's (1953) account of genetic assimilation, in hid case with Drosophila. There is no reason to suppose that the rules are different in species of domestic livestock.

Background genetic variation may be tiresome in terms of identifying major genes, but nevertheless may be useful when it comes to their exploitation in animal production. It may well be possible to alleviate some of the deleterious effects of the gene by selection, as other genes will also contribute to the affected trait. Cockrem's (1959) pioneering work on "breaking" a genetic correlation was developed by Eisen (1978), who discussed selection methods for enhancing the positive aspects of antagonistic traits. There is no reason in principle why similar methods should not be applied to the manifold effects of a single gene, to mitigate some of its deleterious effects. Whether this becomes worthwhile or not will depend of course on the potential gains in production to be made from the major gene. It could well be useful to extend studies of such a gene to its mode of action, in physiological or biochemical terms. Not only would this extend our understanding of its pleitropic effects but also perhaps suggest the most appropriate method of reducing its less

desirable consequences. Its exploitation might then be based on rational procedures and rely less on empirical observations.

In future, as we get closer to the underlying physiological and biochemical basis of quantitative traits, individual loci with major effects may become more important. This has occurred in other branches of genetics especially human genetics, where many anomalies are being partitioned into discrete genetic entities. Should this occur with quantitative traits in farm animals, the knowledge can be quickly used because the theory and practical methods of application are available for their exploitation.

SUMMARY

Identified loci with large effects have not been important in the genetic improvement of farm livestock in the past. However, should major genes be identified, methods ara vailable for their exploitation, with some gains in the rates of selection response. With overdominant loci or loci with deleterious effects, special breeding methods or mating strategies may be required. Screening large and diverse populations for genetic variants has been successful in plants, and is suggested also for farm animals. However, large genetic effects are required to detect whether major genes exist, and the methods are not very powerful unless the effects are large. In the exploitation of a single gene, it is important to know all the economic effects of a major gene, or else the programme may be misdirected. Difficulties arise if the gene is affected by the genetic background or by the environment, but these also offer opportunities in modifying the effects of the gene and in its exploitation.

RESUMEN

Los loci con amplios efectos identificados, no han sido importantes para el progreso genético del ganado en el pasado. Sin embargo, para identificar los genes principales, existen métodos para su explotación con algunos avances en los índices de respuesta de selección. Para loci superdominantes, o de efectos nocivos, podrán ser necesarios métodos especiales de cría o estrategias de cruce. La exploración de poblaciones amplias y distintas en sus variantes genéticas ha dado buen resultado en las plantas, y se recomienda también para los animales domésticos. Se necesitan, sin embargo, amplios efectos genéticos para determinar si existen genes principales, y los métodos no son muy poderosos, salvo que los efectos sean profundos. En la explotación de un gen aislado, es importante conocer todos los efectos económicos de un gen principal, pues de lo contrario el programa puede quedar desorientado. Las dificultades aparecen cuando el gen se ve afectado por los antecedentes genéticos o por el ambiente, pero estos ofrecen también oportunidades de modificar los efectos del gen y en su explotación.

REFERENCES

BODMER, W.F. (1960). Interaction of modifiers: the effects of pallid and fidget on polydactyly in the mouse. Heredity, 14: 445-448.

BREESE, E.L. & MATHER, K. (1957). The organisation of polygenic activity within a chromosome in Drosophila. I. Hair characters. Heredity, 11: 373-395.

CLAYTON, G.A., MORRIS, J.A. & ROBERTSON, A. (1957). An experimental check on quantitative genetical theory. I. Short-term responses to selection. J. Genet, 55: 131-151.

COCKREM, F. (1959). Selection for relationships opposite to those predicted by the genetic correlation between two traits in the house mouse (Mus musculus). Nature (Lond.) 183: 342-343.

DAVIE, A.M. (1979). The singles method for segregation analysis under incomplete ascertainment. Ann. Hum. Genet, 42: 507-512.

EIKELENBOOM, F. & MINKEMA, D. (1974). Prediction of pale, soft exudative muscle with a non-lethal test for the halothane induced porcine malignant hyperthermia syndrome. Tijdscr. Diergeneesk, 99: 421-426.

EISEN, E.J. (1978). Single-trait and antagonistic index selection for litter size and body weight in mice. Genetics, 88: 781-811.

FALCONER, D.S. (1981). Introduction to Quantitative Genetics. 2nd ed. London and New York: Longman.

FISHER, R.A. (1950). Polydactyly in mice. Nature (Lond.) 165: 407.

GELDERMANN, H. (1975). Investigations on inheritance of quantitative characters in animals by gene markers. 1. Methods. Theor. Appl. Genet, 46: 319-330.

GO, R.C.P., ELSTON, R.C. & KAPLAN, E.B. (1978). Efficiency and robustness of pedigree segregation analysis. Amer. J. Hum. Genet, 30: 28-37.

HASEMAN, J.K. & ELSTON, R.C. (1972): The investigation of linkage between a quantitative trait and a marker locus. Behav. Genet, 2: 3-19.

HILL, A.P. (1975). Quantitative linkage: a statistical procedure for its detection and estimation. Ann. Human Genet, 38: 439-449.

JAYAKAR, S.D. (1970). On the detection and estimation of linkage between a locus influencing a quantitative trait and a marker locus. Biometrics, 26: 451-466.

KARLIN, S., CARNELLI, D. & WILLIAMS, R. (1979). Index measures for assessing the mode of inheritance of continuously distributed traits. 1. Theory and justifications. Theor. Pop. Biol., 16: 81-106.

LAUVERGNE, J.J. (1968). Catalogue des anomalies hereditaires des bovins. Bulletin technique du Depart de Genetique Animale. CRNZ, Jouy-en-Josas.

MERAT, P. & RICARD, F.H. (1974). Etude d'un gene de nanisme lié au sexe chez la poule: importance de l'etat d'engraissement et gain de poids chez l'adulte. Ann. Genet. Sel. anim., 6: 211-217.

MERRY, A., ROGER, J.H. & CURNOW, R.M. (1979). A two locus model for the inheritance of a familial disease. Ann. Human Genet, 43: 71-80.

MORTON, N.E. & MacLEAN, C.J. (1974). Analysis of family resemblance. III. Complex segregation of quantitative traits. Amer. J. Human Genet. 26: 489-503.

NEIMAN-SØRENSEN, & ROBERTSON, A. (1961). The association between blood groups and several production characters in three Danish cattle breeds. Acta. Agric. Scand., 11: 163-196.

PIPER, L.R. (1972). The isolation of genes underlying continuous variation. Ph.D. thesis, Univ. of Edinburgh.

RAO, D.C., MORTON, N.E., LALOUEL, J.M. & LEW, R. (1979). Analysis of family resemblance. II. Path and segregation analysis of attribute data. Pop. Genet. Lab. Tech. Report No. 7, University of Hawaii, Honolulu.

ROBERTSON, A. (1973). Linkage between marker loci and those affecting a quantitative trait. Behav. Genet., 3: 389-391.

SHRIMPTON, A.E. (1981). The isolation of polygenic factors controlling bristle score in Drosophila melanogaster. Ph.D. Thesis, University of Edinburgh.

SMITH, C. (1967). Improvement of metric traits through specific genetic loci. Anim. Prod., 9: 349-358.

SMITH, C. & BAMPTON, P.R. (1976). Inheritance of reaction to halothane anaesthesia in pigs. Genet. Res. (Camb). 29: 287-292.

SMITH, C. & WEBB, A.J. (1981). Effects of major genes on animal breeding strategies. Z. Tierzücht. Züchtungsbiol. 98: 161-169.

STEWART, J. (1969). Biometrical genetics with one or two loci. The choice of a specific genetic model. Heredity, 24: 211-224.

SOMES, R.G. (1978). Registry of poultry genetic stocks. Bulletin 446, Storr Agricultural Experiment Station Publications, University of Connecticut.

THODAY, J.M. (1961). Location of polygenes. Nature (Lond.), 191: 368-370.

WADDINGTON, C.H. (1953). Genetic assimilation of an acquired character. Evolution, 7: 118-126.

WEBB, A.J. & JORDAN, C.H.C. (1979). The halothane test in genetic improvement programmes: experiments with Pietrain/Hampshire pigs. Acta. Agric. Scand. (Suppl). 21: 418-426.

WEHRHAHN, C. & ALLARD, R.W. (1965). The detection and measurement of effects of individual genes involved in the inheritance of a quantitative character in wheat. Genetics, 51: 109-119.

YOUNG, G.B. (1967). Heredity diseases in livestock. Proc. 85th Annual Congr. of Brit. Vet. Association, Southport.