Studies of Quantitative Genetic Variation

by

Peter D. Keightley

Thesis presented for the degree of Doctor of Philosophy

University of Edinburgh

1988

Declaration

I declare that this thesis has been composed by myself and that any help received in its preparation has been acknowledged.

Acknowledgements

I wish to thank Professor W.G. Hill for his support, encouragement and guidance and for many hours of stimulating discussions. I thank Dr. H. Kacser for his continued interest and especially for encouraging me to return to Genetics for which I am very grateful. I thank Naomi Wray for commenting on a substantial part of the manuscript and Susan Brotherstone for advice which she gave me on numerous occasions. I acknowledge financial support from the Agricultural and Food Research Council.

A large part of this thesis is based on joint papers with Professor Hill (Ch. 3–5 and 8) and with Dr. Kacser (Ch. 7), published or in press. I specifically acknowledge Professor Hill's contribution of the following analysis: infinitesimal model, subsection 3.2.1; variance of genetic variance, subsection 5.3.2; proof of dominance identity, Appendix to Ch. 7.

Table of Contents

ABSTRACT

Models of genetic variation in quantitative traits are investigated. One type of analysis considers the genetic variance maintained with joint effects of mutation, selection, linkage and drift. The model of the mutation process assumes that there is a large number of sites capable of affecting the value of the trait, but there is a distribution of effects of new mutants at these sites. The effects of changes in the parameters of this distribution on genetic variance maintained under artificial and natural selection are investigated. In artificial populations, mutation can contribute substantially to genetic variance maintained and to long-term selection responses. In natural populations, which are commonly observed to have substantial heritabilities for many quantitative traits, a balance between mutation and drift can maintain sufficient variation, but the strength and mode of natural selection is a critical parameter on which we have little information. The effects of linkage are small in both artificial and natural situations especially for species with many chromosomes.

A second type of model of a quantitative character considers the genetic variation of a particular type of quantitative trait, metabolic flux, and the effects of enzyme activity variation on it. It is assumed that enzyme activity variants segregate in a population, and equations are derived for genetic components of flux variation. Both dominance and epistasis are present, but these are small components of variation relative to the additive variance unless allelic effects are very large. Models of pleiotropic gene action are considered.

A general discussion of the likely contribution of molecular techniques to understanding quantitative genetic variation is given. It is argued that molecular techniques involving insertional mutagenesis will lead to a higher level of understanding of quantitative variation at the level of gene action. The problem of the maintenance of variation in quantitative traits, however, has similarities to the problem of the maintenance of molecular polymorphism, for which there has been no clear resolution.

Table of Symbols

Symbols used at one place only in the thesis are defined where they occur. One or two symbols have different meanings in the two parts of the thesis, and this is indicated below.

α	Ch. 2–5. Scale parameter of gamma distribution.
α	Ch. 6. Average effect of a gene substitution.
β	Shape parameter of gamma distribution.
ε	Ch. 2–5. Root expected squared value of mutation effect.
ε	Ch. 6. Interaction term.
λ	Genomic mutation rate.
μ	Locus mutation rate.
σ	Phenotypic standard deviation.
Φ	$N_e a^2/[8(w^2 + \sigma^2)] = N_e s^*$.
a	Ch. 2–5. Difference between homozygotes of effect of allele.
a	Ch. 6–7. Group enzyme activity.
b	Group enzyme activity.
c	Group enzyme activity.
D	Dominance index.
$\sum D$	Disequilibrium covariance component.
E	Enzyme name.
E	Enzyme activity.
H	Heterozygosity.
h^2	Heritability.
i	Selection intensity.
J	Flux.
K	(Unsubscripted). Group enzyme activity of 'constant' enzymes.
$K_{i,j}$	Equilibrium constant.
k	Factor for strength of selection.
L	Length of chromosome in morgans.
M_i	K_m = Michaelis constant.
N	Parental population size.
N_1	Effective population number in stabilizing selection phase.
N_2	Effective population number in directional selection phase.

N	Number of individuals measured to estimate genetic variance.
N_e	Effective population number.
n	Number of loci.
P	Proportion of mutant effects (or distribution) positive.
q	Frequency of wild type allele.
R	Response to selection.
s	Selective advantage of homozygote with mutant alleles over wild type.
s_1^*	Selection coefficient in stabilizing selection phase.
s_2	Selection coefficient in directional selection phase.
S_i	Substrate name i
S_i	Substrate concentration i
S	Common substrate in branched pathway.
t	Generation number.
T	Progeny population size.
V_A	Additive genetic variance.
V_D	Dominance variance.
V_E	Environmental variance.
V_g	Genic variance.
V_I	Interaction variance.
V_M	New mutational variance arising per generation.
V_T	Total genetic variance.
V_i	V_{MAX} = maximal velocity.
v	Rate of reaction.
$W()$	Fitness function.
w	Inverse measure of strength of stabilizing selection.
X	Ch. 2–5. Phenotypic value of an individual.
X_A	Source substrate.
X_B	Sink substrate.
X_C	Sink substrate.

Chapter 1
General Introduction

1.1. Ubiquitous quantitative variation

The scientific study of the inheritance of continuous characters originated soon after the publication of Darwin's "Origin of Species". The rediscovery of Mendel's paper was followed by a debate between 'Mendelians' and 'Biometricians' on the connection between Mendel's discoveries and the inheritance of quantitative traits. The arguments were resolved in a paper by Fisher (1918) who published correlations between relatives on the assumption of Mendelian inheritance. (Earlier papers on this subject by Weinberg (1909a,b, 1910) were ignored). Subsequently, the foundations of quantitative genetics theory were laid with major contributions from Wright (1921) and Haldane (summarized in 1932). Provine (1971) gives a lucid account of the development of the theoretical basis of quantitative genetics.

Variation in quantitative traits is important because selective improvement of domestic species depends on it, and because it is generally held to be the 'raw material' for evolutionary changes. There is ample evidence for the near universality of genetic variation in quantitative traits. In two recent studies, Mousseau and Roff (1987) and Roff and Mousseau (1987) summarized published values of narrow sense heritabilities of many characters from natural populations of diverse animal species, the latter paper devoted to the Drosophila genus. For morphological characters the average heritability was 46% for animals in general and 32% for Drosophila. Heritabilities of life history and physiological traits were generally lower. For traits closely connected with fitness there was usually a small but detectable additive genetic component of variation. Caution is needed in interpreting such data because various environmental factors can cause heritability estimates to be biased upwards (Falconer, 1981 Ch. 9). However, such evidence together with the general observation of responses to artificial selection (Lewontin, 1974 Ch. 2) leads to the inevitable conclusion that substantial additive genetic variation is present in most characters in artificial and natural populations. The presence of such ubiquitous genetic variation raises two important problems. The first is its maintenance. The observation of heritable variation of any character is the

result of past influences on the population of the major evolutionary forces of drift, migration, mutation and selection. There are several important questions which relate to the maintenance of genetic variation: How important is finite population size relative to the other forces? What is the nature of the selective forces and do these maintain or destroy variation? How quickly does mutation generate quantitative variation? What are the implications of spatial and temporal heterogeneity of the environment? The second problem is that of the nature of quantitative variation. One aspect is the identification of the genes responsible for the genetic variation in the character and the elucidation of their biological functions. A second is understanding the properties of alleles contributing to genetic variation in the population, the distributions of allelic effects and frequencies and their dominance and epistatic relations.

This thesis examines aspects of these two problems theoretically. It falls into two distinct parts. The first part (Ch. 2–5) concerns the role of mutations in contributing to selection responses and to genetic variation in finite populations. In the second part (Ch. 6–7), specific models of the action of a class of gene products, the enzymes, are investigated in relation to variation in quantitative characters.

1.2. Mutation and quantitative variation

Mutation is the primary source of genetic variation in quantitative characters. Mather and Wigan (1942) performed the first experimental investigation of the influence of new mutations on a quantitative trait by measuring the rate of change of bristle score from selection of inbred sublines of *Drosophila melanogaster*. Clayton and Robertson (1955) performed similar experiments and developed theory on the contribution of mutations to variation in quantitative traits. The important parameter of new mutational variance arising per generation, V_M, was introduced. V_M is usually expressed as V_M/V_E, the new genetic variance as a proportion of the environmental variance of a (usually) inbred line. Clayton and Robertson also gave the important result that for additive genes in the absence of selection the expected variance in a finite population at equilibrium is $2N_e V_M$, with N_e the effective population number. Their estimate of V_M for Drosophila abdominal bristles was $10^{-4} V_E$, that of Mather and Wigan (1942) was $10^{-3} V_E$ which was suggested as an 'upper limit'. Thus the contribution from mutation per generation to the variation in the character is relatively small. Experiments designed to measure V_M/V_E are

difficult because they require many generations and need to be large in view of the rarity of spontaneous mutations. Some estimates for various characters of mice come from measurements of the divergence between sublines of inbred strains which have been maintained for many generations ostensibly for other purposes (Bailey, 1959; Grewal, 1962; Festing, 1973). Lynch (1988) has extensively reviewed published estimates of V_M/V_E for diverse characters in a number of animal and plant species. Where more than one estimate is available, there is usually substantial variation among estimates. The 'upper limit' figure of 10^{-3} suggested by Clayton and Robertson (1955) for Drosophila bristle score is, however, probably a reasonable one for the spontaneous rate because there is a fair degree of agreement among estimates from various methods. The estimates from the divergence rates of mouse inbred sublines are generally higher. It would not be surprising to find widely differing rates of generation of mutational variation between species because factors such as varying generation times and numbers of cell divisions per generation of the germ line are likely to influence spontaneous mutation rates (Nei, 1987 Ch.3). An additional source of variation may come from variation in the efficiency of DNA repair systems. Further, the distribution of effects of new mutant alleles, relating the expected number of mutational effects in different size classes, is likely to vary widely between characters. In Drosophila, the discovery that quantitative variation can be generated by mobilisation of transposable elements into new genomic sites (Mackay, 1985) also suggests that the input of variation from mutation can differ substantially between populations because the induction of transposition depends on specific dysgenic crosses. A similar phenomenon also occurs in a mammalian species, the mouse, in which crosses of specific inbred lines lead to mutations due to the mobilisation of proviruses (Jenkins and Copeland, 1985).

1.3. Selection, mutation and variation in quantitative characters

Recently, there has been renewed interest in the contribution of mutation to responses in artificial populations stimulated by the observation of long continued responses in some experimental populations (e.g. Dudley, 1977; Enfield, 1980; Yoo, 1980), and in commercial populations (Smith, 1984). In natural populations, the maintenance of variation in quantitative characters has received much attention recently because of the conflict between the results of theoretical studies on the maintenance of genetic variation by mutation-stabilizing selection balance by Lande (1976) and Turelli (1984).

Interest has also been stimulated by the discovery of vast quantities of variation at the molecular level by molecular techniques starting with starch gel electrophoresis of enzyme variants (Harris, 1966; Lewontin and Hubby, 1966) and presently by analysis of DNA by various methods.

1.3.1. Mutation in artificially selected populations

Hill (1982a,b) considered the contribution of new mutations to artificial selection response rates and concluded that mutation can be important within the time scales of commercial selection programmes and in long-term selection experiments in spite of the relatively small contribution of new mutations to the genetic variance each generation. The model involved mutations affecting the trait from the whole genome rather than a specific set of loci 'controlling' the trait. The relevant parameters of the mutation process reduced to the expected number of mutations per haploid genome per generation and the shape and scale of the distribution from which it was assumed mutational effects are independently sampled. The shape and scale of the distribution were shown to have strong influences on the expected response from new mutations in the early generations after the commencement of selection. Subsequently, Hill and Rasbash (1986) showed that the variance of response between lines also strongly depends on the shape and scale of the mutational distribution. For the case of additive genes, however, the asymptotic response rate from fixation of mutant alleles is proportional to the product of strength of selection and the asymptotic variance, namely $2N_e V_M$ as in the neutral case, with little dependence on the scale and shape of the mutational distribution. The quantitative predictions are therefore heavily dependent on the magnitude of V_M/V_E as discussed above, but an important conclusion was that the asymptotic response is proportional to the effective population size. Further analysis by Keightley and Hill (1983) of the effects of linkage using a model of a chromosome with an infinite number of independently mutable sites led to the conclusion that linkage has little influence on asymptotic response rates unless it is very tight.

1.3.2. Mutation and stabilizing selection in natural populations

The optimum model of stabilizing selection is a popular model of natural selection of quantitative characters because it predicts a stable value of the character. The character has a fixed optimum value and deviation in either

direction implies a reduction in fitness. Of course, this is only one possible model of natural selection, but it has been subject to a great deal of analysis. The following is a summary of some of the most important papers in the area.

1. Latter (1960) gave the first important analysis of the equilibrium behaviour of a mutation-stabilizing selection balance. The 'nor-optimal' model of stabilizing selection (Haldane, 1954) was analysed with a model of a locus with up to two alleles segregating in an infinite population. A simple formula was derived for the genetic variance at equilibrium. Based on arguments concerning per locus mutation rates, numbers of loci thought to 'control' quantitative traits and likely strength of stabilizing selection in nature, Latter concluded that mechanisms other than a balance between mutation and selection were necessary to explain observed natural heritabilities.

2. Kimura (1965) analysed the equilibrium behaviour of a locus with multiple alleles of small step-wise effects segregating in an infinite population. An expression was derived for the equilibrium genetic variance under mutation-stabilizing selection balance which was quite different from that of Latter (1960). Kimura made the important conclusion that the model predicts that the equilibrium distribution of allelic effects segregating at the locus is normal. Kimura did not explicitly consider whether the model could account for natural heritabilities.

3. Latter (1970) extended Kimura's (1965) analysis and incorporated finite population size. Interestingly, Latter used a fundamentally different model from his 1960 paper (described above) which is perhaps evidence for the influence of Kimura's (1965) results.

4. Bulmer (1972) analysed mutation-stabilizing selection balance for a two allele model which allowed for the possibility of large differences in effect between the alleles, and by incorporating diffusion theory developed by Kimura (1964), obtained formulae for the variance maintained in finite populations. Bulmer showed that less variation is maintained in a finite population than predicted for the absence of drift, but did not discuss the variation which might be maintained with parameters 'appropriate' to natural populations.

5. The paper of Lande (1976) was important for arguing that high heritabilities can be explained by a mutation-stabilizing selection balance without need of other explanations. Lande attempted to fit experimentally determined estimates

of V_M and number of 'effective factors' influencing a quantitative trait (Falconer, 1981), to a model of an infinite population with assumptions based on results of Kimura (1965) (see above). Lande concluded that high heritabilities can be maintained even for strong stabilizing selection with values of the mutation parameters which were perceived to be reasonable.

6. Turelli (1984) also extensively reviewed experimental estimates of the important parameters, but using a different model from Lande (1976), concluded that the question of the maintenance of natural heritabilities depends on the values of parameters for which there is insufficient data. Turelli showed that Kimura's (1965) result that the distribution of allelic effects segregating at the loci contributing to the trait is normal depends on high mutation rates per locus and small selective values of new mutant alleles. In Turelli's "House of Cards" model the effects of new mutant alleles were assumed to be sufficiently large to swamp existing variation at the loci, and Turelli argued that such a situation is likely in nature given the values of experimentally measured parameters such as per locus mutation rates and V_M/V_E for quantitative traits. The formula obtained for the equilibrium genetic variance was the same as that derived by Latter (1960) and Bulmer (1972).

These issues are still being discussed in the literature (e.g. Bulmer, 1988), and they will be examined in more detail later in this thesis (Ch. 4–5).

1.3.3. Other models of the maintenance of variation

The popularity of the single character model of stabilizing selection is in part due to its tractability. Models such as frequency-dependent selection (Clarke and O'Donald, 1964), and density-dependent selection (Clarke, 1972) can lead to the maintenance of variation through marginal overdominance of the heterozygote with respect to fitness, as can temporal and spatial variation of the environment. Mackay (1981) showed that significantly greater genetic variation was maintained in various traits of Drosophila in populations subject to environmental heterogeneity. The above mechanisms, however, are beyond the scope of this thesis. Other models are considered, however.

Neutrality. Clayton and Robertson's (1955) result that the equilibrium variance for additive genes in a mutation-drift balance is $2N_e V_M$ is important because it shows that without selection, mutation is an adequate force to explain

observed natural heritabilities. So for example, with V_M/V_E of 10^{-3} and an effective population size of 10^3, the expected heritability is 2/3. The time taken to reach the equilibrium in large populations can be very long (Chakraboty and Nei, 1982).

Lynch and Hill (1986) performed an extensive analysis of mutation-drift balance and included dominance and linkage in their models, which were shown to have little influence on the expected variance, but to have an important influence on the variance of the genetic variance between independent lines. The model analysed was step-wise in the sense that the effect of each new mutant was assumed to be added on to the value of the allele already present at the locus. The genetic variance can therefore increase without bounds with increasing N_e, as can the asymptotic response to directional selection (Hill, 1982a,b). Cockerham and Tachida (1987) analysed similar problems, but assumed a model formally the same as the "House of Cards" (cf. Turelli, 1984), i.e. the effect of a mutant allele *replaces* the current allelic state. This gives different behaviour from the step-wise model described above because the variance with neutral genes is finite in an infinite population and similarly there can be a directional selection limit. Clearly, the effects of mutations in natural situations vary between the step-wise and House of Cards extremes.

Pleiotropy. An important extension to the stabilizing selection model is the incorporation of the effects of pleiotropy. Two notable studies by Lande (1980) and Turelli (1985) extended Kimura's (1965) and Turelli's (1984) analyses respectively, and assumed stabilizing selection simultaneously acting on more than one character with new mutations simultaneously affecting all characters. As with the single trait analysis described above, the two models behave quite differently with Turelli's House of Cards assumptions predicting a large impact of pleiotropic selection on the variance maintained in the character.

It is unlikely that pleiotropic effects occur solely through characters with intermediate optima, as the effect of pleiotropy is often likely to be through characters closely connected with fitness itself. For this reason, the single character optimum model has been frequently criticised in the literature (e.g. Robertson, 1973; Falconer, 1981 Ch. 20; Gillespie, 1984). Hill and Keightley (1988) considered the variance maintained in a trait with directional selection where mutations have effects on the trait and correlated pleiotropic directional effects on fitness. As expected, the presence of such pleiotropy leads to a

reduction in the variance maintained in the character, but also has relatively more impact on the selection response. The extent depends on genetic parameters such as the variance of the distribution of effects of mutant alleles and the correlation between the effects of mutants on the trait and fitness. Such an analysis would also be appropriate for natural populations.

Heterozygote Superiority. Intuitively, heterozygote superiority for fitness is an attractive candidate for explaining the maintenance of genetic variation in the presence of other forces which tend to reduce it. It has been argued, however, that heterozygote advantage at the level of the 'primary effect' of the gene is unlikely, because the heterozygote is usually intermediate if it is at an enzyme activity locus (Kacser and Burns, 1981). Heterozygote advantage for fitness is more likely to be due to opposing directional pleiotropic effects on its 'major components' (Rose, 1982). This can also lead to apparent intermediate optima for the individual characters when regressed on fitness (Hill and Keightley, 1988). Frequently quoted examples of loci showing heterozygote superiority for fitness are usually examples of this phenomenon. For example the advantage of the heterozygotes for the thalassaemia alleles is due to negative pleiotropic effects on efficiency of oxygen transport of erythrocytes opposed by a positive influence on resistance to malaria (Weatherall and Clegg, 1981). Gillespie (1984) analysed a model of an infinite population and showed that substantial variation can be maintained from alleles of additive effect on the trait if there is heterozygote superiority in fitness. There is some doubt, however, whether heterozygote superiority (whatever the mechanism) can be a general explanation for the maintenance of quantitative variation because such superiority can actually accelerate the rate of fixation at a locus in small populations if the stable gene frequency is near either of the extremes (Robertson, 1962).

1.4. The effects of linkage

Quantitative traits are affected by alleles segregating at many loci which may or may not be physically linked on the same chromosome. Selection tends to generate non-random combinations of alleles at different loci, and recombination tends to break them down. Such non-randomness is termed disequilibrium. The term *disequilibrium* rather than *linkage disequilibrium* is generally used here since disequilibrium can be generated in the absence of linkage. In quantitative genetics theory linkage is important because of its

influence on selection responses and on the expectation and variances of genetic variances in natural and artificial populations.

1.4.1. Linkage and quantitative genetic variation

Disequilibrium in a population can be measured as a covariance component, a function of values and frequencies of alleles and frequencies of haplotypes of pairs of loci, which may be either positive or negative. An excess of coupling haplotypes generated by directional or stabilizing selection leads to an expectation of less additive variance compared to random mating and to the presence of a negative disequilibrium covariance component. As a consequence, selection responses predicted from the additive genetic variance in an unselected population are reduced in second and subsequent generations of selection (Bulmer, 1980 Ch. 9). This is a general qualitative prediction for which there is evidence from the results of selection experiments (Atkins and Thompson, 1986). If drift is the only force tending to change gene frequencies, the expected disequilibrium in the population is zero, but the disequilibrium varies stochastically about zero due to random changes in gene frequency. Linkage has the theoretical consequence of increasing the variance of the variance of genotypic values between sublines (Bulmer, 1976; Avery and Hill, 1977).

Recombination reduces disequilibrium by randomly generating all genotypes irrespective of the initial conditions. With free recombination, the disequilibrium in the progeny is reduced on average by a factor of one-half each generation compared to the parents. The effect of a finite amount of recombination is also to break down disequilibrium, but to a lesser extent than free recombination depending on details of the recombination mechanism. For example, the presence of crossover interference is similar to the behaviour in its absence, but with a reduced chromosome length (Avery and Hill, 1979).

In the presence of mutation and selection the effects of linkage on equilibrium variances and responses depend, in theory, on assumptions of the mutation process. Lande (1976) investigated a multi-allele model of a finite number of loci and mutations of tiny effect, and concluded that linkage has a small influence on the variance maintained with a mutation-stabilizing selection balance. Using a different model, with finite effects of mutant alleles occurring at an infinite number of possible sites, Keightley and Hill (1983) analysed

linkage in finite populations with directional selection and showed that one crossover or more per chromosome per generation is very similar to free recombination.

Theoretical analysis shows that linkage tends to retard selection responses, but an apparent paradox arises because in the case of multiplicative gene action, selection does not immediately generate disequilibrium (Felsenstein, 1965). Hill and Robertson (1966) used Monte Carlo simulation of a two locus model to compare additive and multiplicative models of gene action and showed that selection responses are retarded to a similar extent in both situations. The result was explained in terms of the effective population size in which gene frequency changes take place at each locus, which is reduced by linkage. Keightley and Hill (1983) also observed little difference between additive and multiplicative models.

Robertson (1970, 1977) simulated the effects of linkage on responses to directional selection using models of discrete loci and infinitesimal effects spread evenly along a chromosome respectively. Robertson attempted to describe the behaviour of the models in terms of as few parameters as possible and concluded that in general linkage needs to be very tight to have substantial effects either on response rates or limits.

1.5. Metabolic models of quantitative characters

Usually, the biological mode of action of the genes affecting quantitative characters is not or cannot be considered. This stems from the present lack of understanding of the relationship of most characters to the genetic system. One class of genes, those controlling enzyme activities, is substantially better understood than other classes (for example, genes controlling development).

Enzymes operate within the metabolic system which is represented graphically as the 'metabolic map'. The map shows the arrangement of the enzyme catalysed reactions in complex pathways. The kinetics of the system depend on, among other parameters, enzyme activities which are under genetic control. Since there is experimental evidence that enzyme activity is polygenically determined (Paigen, 1979; Laurie-Ahlberg *et al*, 1980), enzyme activity itself can therefore be regarded as the lowest level of quantitative character measurable in metabolism. At the next level up, concentrations of metabolic intermediates

and metabolic fluxes are also quantitative characters. Because of the interconnected nature of the system, any flux or metabolite pool is necessarily a function of many, perhaps all, enzyme species.

The extent to which changes in enzyme activity influence fluxes and metabolite pools is the subject of studies of the quantitative control of metabolism (reviewed by Kacser and Porteous, 1987). The concept of 'rate limiting step' is replaced by 'control coefficient' which describes the fractional change in flux resulting from a small fractional change in enzyme activity. Most studies show that the control of flux and metabolite pool is shared among many enzymes and cannot be uniquely assigned to any one (e.g. Flint, Porteous and Kacser, 1980; Groen *et al.*, 1982; Salter, Knowles and Pogson, 1986; Woodrow, 1986; Dykhuizen, Dean and Hartl, 1987). Theory of control coefficients is applicable to small changes in enzyme activity, but mutations can cause large changes in enzyme activity and the effect on fluxes and pools is not fully described by theory developed for infinitesimal changes.

Kacser and Burns (1981) modelled the effects of finite changes of enzyme activity on the dominance relations of the phenotype flux in diploids. Using results obtained from a general control analysis of metabolic systems and specific models of simple pathways, the relationship between flux and enzyme activity was shown in general to be hyperbolic, and with constraints imposed by the distribution of control coefficients, most enzyme activities in any pathway lie on the plateau of the curve (Fig.1.1). Thus a null mutant is most likely to be recessive for the phenotype flux, while a small change in enzyme activity gives a nearly intermediate flux value of the heterozygote. Kacser and Burns argued that many characters are closely related to metabolic flux and that null mutations in such characters tend therefore to be recessive. Thus, dominance is a consequence of the constraints imposed by the metabolic system without needing a direct evolutionary explanation of modifiers (Fisher, 1928). On the other hand, small changes in enzyme activity are likely to be nearly additive.

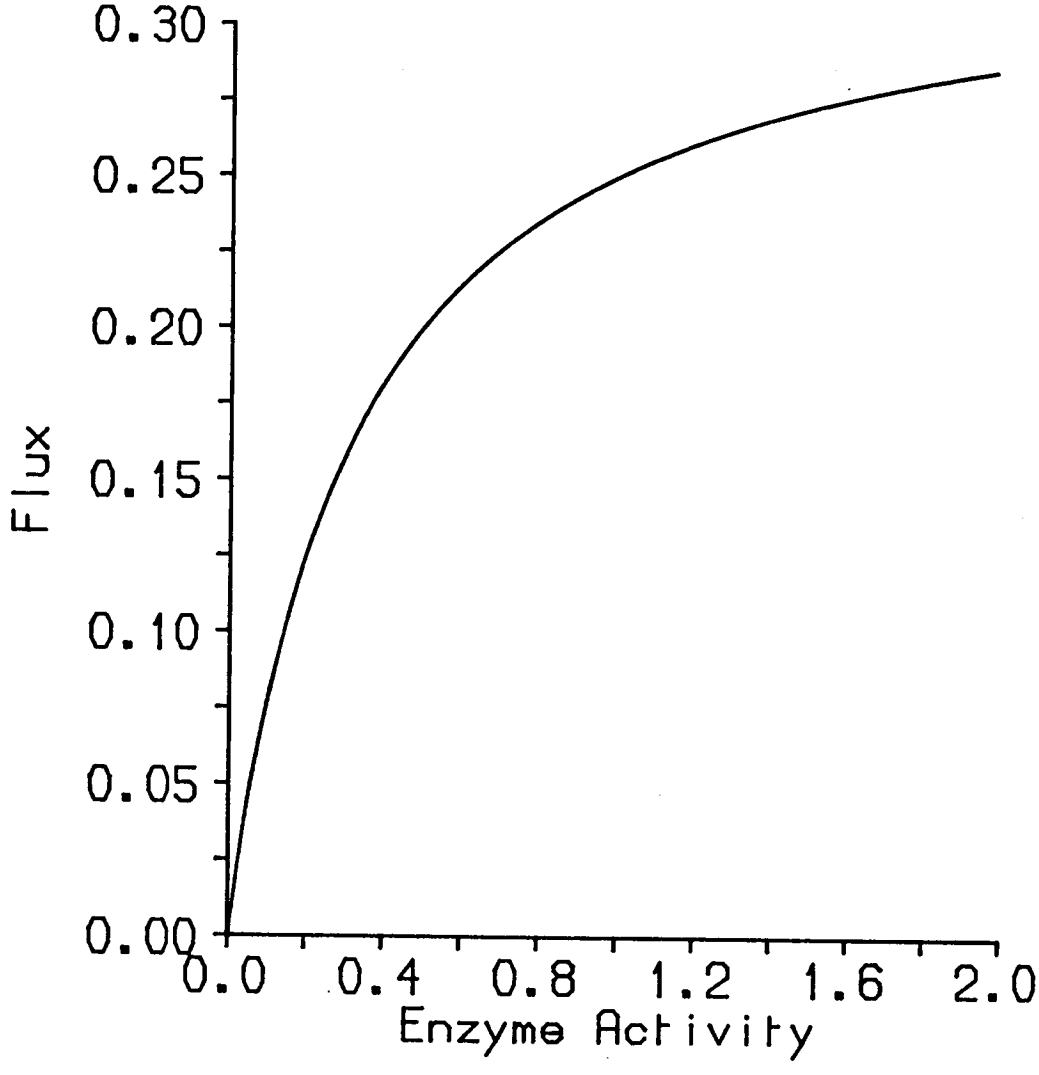


Fig.1.1. The response of flux to change in one enzyme activity in a linear pathway at steady state of four unsaturated enzymes of unit activity (see Kacser and Burns 1973, 1981, and Ch. 6 for details of the model).

Non-additive allelic interactions within loci are an inevitable consequence of the kinetics of metabolic pathways. Similarly, epistasis or interaction between alleles at different loci is also generated in metabolic systems. The following simple example illustrates the point. Consider a null mutant in a metabolic pathway. The flux in the pathway is zero because the pathway is blocked. The presence of another null mutation at a different step is epistatic to the first because the flux is also zero with the double mutant. Enzyme activity alleles segregating in a diploid population therefore generate both dominance and interaction variance in the quantitative characters metabolic flux and metabolite concentration. Kinghorn (1987) attempted to analyse the nature of epistatic interactions for growth and reproductive traits in guinea pig data accumulated by Wright (1922). Although Kinghorn concluded that the data best fitted a model of additive-by-additive interactions, the model of epistasis was not based on modern ideas on the kinetics of multi-enzyme systems (Kacser and Burns, 1973; Heinrich and Rapoport, 1974). A further feature of the effects of enzyme variation on fluxes and pools is pleiotropy. This occurs because different fluxes are coupled to one another by stoichiometry or by competition for common substrates at branch points. It is intuitively obvious that fluxes widely separated on the 'metabolic map' are almost independent of one another with respect to any enzyme variation and this is why mutants have specific and characteristic effects. The directions of pleiotropic coupling and extent to which different fluxes are related are not intuitively obvious, however.

Dykhuizen, Dean and Hartl (1987) (see also Hartl, Dykhuizen and Dean, 1985) have investigated the fitness consequences of mutations at various sites in the *lac* operon of *Escherichia coli* in lactose limited chemostats. In this environment, fitness is linearly related to flux, and competitive ability of mutants against a tester strain gives a good measure of relative flux. The relationship between enzyme activity and fitness for mutants at both the β-galactosidase and β-galactoside permease genes is hyperbolic, as predicted by metabolic control theory. Further, a model of a 'linear unsaturated pathway' with two variable loci (β-galactosidase and permease) and a 'constant' portion of the pathway fitted the chemostat data very well.

1.6. Subject matter of this thesis

This thesis is a theoretical investigation of two important general problems connected with the nature of quantitative genetic variation. The first concerns the contribution of mutation to variation in quantitative traits in finite populations. The problem of explaining quantitative variation is important because many evolutionary changes at the phenotypic level are brought about by the action of natural selection on quantitative variation, and because selection responses in artificial populations depend on quantitative variation which has arisen by mutation in the base populations or since selection has started. Although others have analysed this problem, few have considered the consequences of assuming a distribution of effects of mutant alleles and the consequences of changes in its parameters on the amount and nature of the variation maintained. Here, these aspects of the mutation process and the interactions of drift, selection and linkage on variance maintained and selection responses are investigated (Ch. 2–5). The second general problem concerns the nature of variation in quantitative traits related to the biochemical system. Most models of quantitative characters do not consider constraints imposed by the underlying biochemical system. Such constraints influence the genetic variation we measure (additive/non-additive), the couplings between characters (pleiotropy) and responses to selection. Here, models of quantitative traits based on biochemical assumptions are investigated (Ch. 6–7).

1.6.1. Outline

Chapter 2 describes the genetic model and the analytical and simulation methods which are used in the three chapters on the contribution of mutation to quantitative variation which follow.

Chapter 3 is an analysis of the interactions of mutation, linkage and drift with responses to directional selection. The results of a Monte Carlo simulation model are compared to the results of an analytical model of tiny gene effects ('infinitesimal model'). The effects of linkage where mutant effects are sampled from a skewed distribution, and in particular, the operation of "Muller's ratchet" are investigated.

Chapter 4 is an investigation of models of quantitative variation maintained at a mutation–stabilizing selection balance for finite populations. Much of the

analysis concentrated on the consequences of distributions of effects of new mutant alleles. The effects of linkage are investigated by Monte Carlo simulation.

Chapter 5 extends the analysis in Chapter 4. Expressions for the variance of the genetic variance between independent lines are derived. Patterns of selection response to directional selection from characters previously under stabilizing selection are simulated and compared to the results of experiments. The magnitude of genetic parameters which underlie selection responses are discussed.

Chapter 6 introduces some simple biochemical models, and uses these to derive expectations of non-additive components of variance in metabolic flux caused by variation in enzyme activities segregating in a population. Implications for the genetics of quantitative variation are discussed.

Chapter 7 is an investigation of dominance of mutants affecting characters pleiotropically coupled by the metabolic system.

Chapter 8 is the general discussion, but concentrates on the consequences of new molecular methods and technology for quantitative genetics.

Chapter 2
Model and Methods

2.1. Introduction

In chapters 3–5, interactions between selection, mutation and linkage are investigated in a model of a finite population. The purpose of this chapter is to describe the assumptions of the biological model, then to describe the methods used to analyse it and the ways in which they relate to each other. Central to the model is the idea of a distribution of effects of mutant alleles. Consider one locus which affects the value of a quantitative character. It is assumed that recurrent mutation can change the value contributed by the locus, that the effect of a new mutant allele is *added* to the current value, and that the effect is sampled from a distribution. Much of the following three chapters is concerned with the effects of changes in the parameters of this distribution (i.e. standard deviation, skewness, kurtosis) on genetic variation and response to selection. The basic aim is to model the processes of mutation, selection, drift and recombination accounting for mutations occurring at all loci in the genome which affect the value of the trait. Three main approximations are used to investigate the model. (i) Analytical. By using diffusion theory (Kimura, 1969), it is possible to derive formulae describing the equilibrium behaviour of a single locus with recurrent mutation. (ii) Transition matrix. A matrix of change of allele frequency at a locus in a finite population is defined. Using this matrix, it is possible to predict the genetic variance contributed by a new mutant allele during its lifetime and its fixation probability. The equilibrium and dynamic behaviour can be predicted. Only single loci can be considered. (iii) Monte Carlo Simulation. This method relies on computer programs using random number generators and allows the effects of linkage and the simultaneous effects of alleles segregating at more than one locus to be assessed. The dynamic and equilibrium behaviour can be measured. In contrast to the previous methods, results have sampling errors.

The following section describes the basic assumptions used in all three approximations. Since the assumptions are the same or very similar, results from each approximation are comparable to one another, and this is done in various places in the results chapters.

2.2. Finite population model

Basic Assumptions. The population is assumed to consist of N monoecious diploid individuals with constant population size, random mating, random family size and non-overlapping generations. Selection is sufficiently strong or the population size sufficiently small that no more than two alleles segregate at any time at each locus contributing to variation in the quantitative character. In some cases, Monte Carlo simulation runs with multiple alleles were used to investigate the consequences of relaxing this restriction. The frequency of the higher valued allele at a locus is q, and the difference in value between the homozygotes is a. Effects are assumed to be additive within and between loci, i.e. there is no dominance or epistasis.

Mutation. It is assumed that the population size and mutation rate are such that a new mutation is unlikely while an existing mutation is segregating, but such a mutation can occur later and its effect is a random variable sampled from a distribution of effects of new mutant alleles, a parameter of the model. The value of the new mutant effect is added to the current value at the locus. The mutation model is therefore step-wise. The expected number of mutations appearing per haploid genome per generation is λ and these occur independently. The increment in variance each generation from mutation is

$$V_M = \lambda E(a^2)/2 \qquad (2.1)$$

(Hill, 1982a). The distribution of mutant effects is time-invariant. For modelling purposes the gamma distribution was chosen since it has a wide range of properties if suitable values are given to its parameters. The density function of mutants having an increasing effect on the trait is given by

$$f(a) = \alpha^{\beta} e^{-\alpha a} a^{\beta-1}/\Gamma(\beta), \qquad (0 < a < \infty) \qquad (2.2)$$

where $\Gamma(.)$ is the gamma function and the moments are $E(a) = \beta/\alpha$, $E(a^2) = \beta(\beta + 1)/\alpha^2$. Examples of the gamma distribution for a range of shape parameter, β, are shown in Fig. 2.1. The parameter α defines the scale of the distribution and β its shape. In practice, the scale was defined by the parameter

$$\varepsilon = [E(a^2)/V_E]^{1/2} = [\beta(\beta + 1)/(\alpha^2 V_E)]^{1/2}, \qquad (2.3)$$

where V_E is the environmental variance. With shape parameter $\beta = 1$, $f(a)$ is an exponential distribution; as $\beta \to 0$ the distribution approaches the geometric distribution and becomes increasingly leptokurtic with a large spike near $a = 0$ and a long tail; with $\beta \to \infty$, the distribution approaches the limiting case of all mutant effects equal. This model was used to investigate the effect on genetic variation and selection responses of different proportions of positive and negative values of effects of mutant alleles. In practice, therefore, mutant effects were sampled from a gamma distribution of scale parameter ε with sign randomly allocated, and probability P of being positive. For the resulting distribution, referred to as the 'reflected gamma distribution', $E(a) = \beta(2P - 1)/\alpha$, and $E(a^2)$ and thus ε are unchanged.

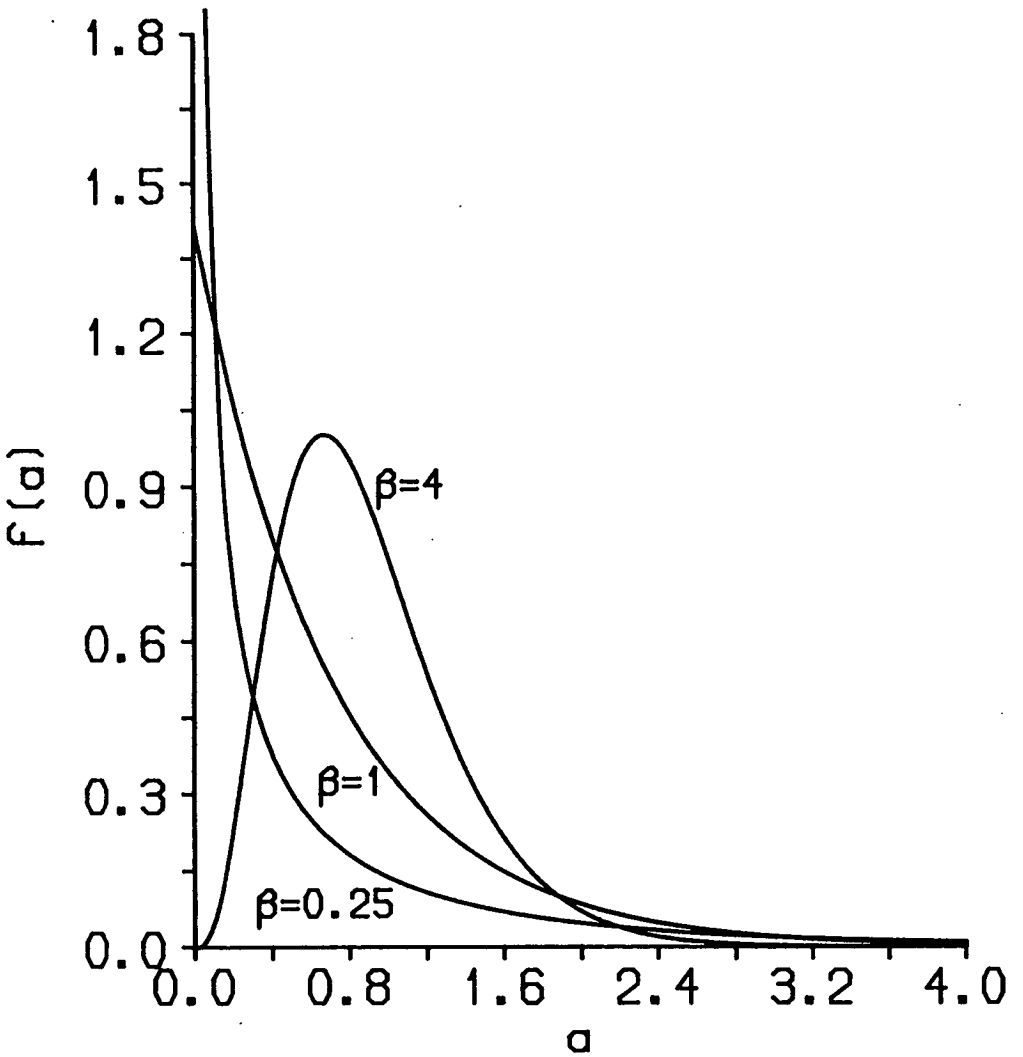


Fig. 2.1. Examples of the gamma distribution, $f(a) = \alpha^{\beta} e^{-\alpha a} a^{\beta-1}/\Gamma(\beta)$ for three values of the shape parameter β. The parameter α describes scale rather than shape and its value is such that $E(a^2) = 1$ for each curve.

2.2.1. Definition of variance components

The phenotypic value of the quantitative character is X, the sum of effects of alleles on each of the two chromosomes, plus a random normally distributed environmental effect of mean zero and variance V_E. The *additive genetic variance*, V_A, of the character is simply the variance of breeding values. The *genic variance* is defined as

$$V_g = \sum a_i^2 q_i(1 - q_i)/2, \tag{2.4}$$

where the summation is over all loci affecting the trait. The analytical and transition matrix methods consider only single loci, so ignore disequilibrium and the effects of selection and drift on it. Only the genic variance and its variance are calculated in this case. The Monte Carlo simulation simultaneously models the fates of all segregating alleles which exist in the population, so the genic and additive variances and therefore the disequilibrium component can be estimated. In the absence of selection, the genic variance at generation t can be obtained from the recurrence

$$V_{g,t+1} = V_{g,t}(1 - 1/2N_e) + V_M, \tag{2.5}$$

where N_e is the effective population number. The genic variance decreases by a factor $1 - 1/2N_e$ due to drift and increases by V_M units from mutation each generation. For additive genes in the absence of selection at a mutation–drift balance, the expected value of the additive variance is the same as that of the genic, namely $2N_e V_M$. Selection can cause the appearance of a disequilibrium component due to the selective advantage or disadvantage of combinations of alleles at different loci which have arisen by chance. Drift generates no net disequilibrium, but in a finite population, disequilibrium varies stochastically due to chance changes in allele and genotype frequencies. This component is given by

$$\sum D = \sum_{i \neq j} \sum a_i a_j D_{ij}/2, \tag{2.6}$$

where $D_{ij} = f_{ij} - q_i q_j$, and f_{ij} is the frequency of the corresponding gamete. $\sum D$ is the sum of the covariances between the values of pairs of loci in the population, and can be either negative or positive. The additive, genic and disequilibrium components are related by,

$$V_A = V_g + \sum D. \tag{2.7}$$

There is also a covariance component due to deviations from Hardy–Weinberg equilibrium (Bulmer, 1976). This is ignored here since it is transient and disappears after one generation irrespective of linkage.

2.2.2. Selection

This subsection describes some approximations for the process of selection used by the analytical and transition matrix methods. Two models of selection are investigated. In Ch. 3, the consequences of directional selection in finite populations are modelled. In Ch. 4, a mutation–stabilizing selection balance is modelled. In Ch. 5, the interactions between the two modes of selection are modelled.

Directional selection. Let s be the selective advantage of the homozygote carrying the mutant alleles over the 'wild type' and $s/2$ be the selective advantage of the heterozygote. The change in gene frequency from one generation of selection is approximated by,

$$\Delta q = sq(1 - q)/[2(1 + sq)] \tag{2.8}$$

(Falconer, 1981). With truncation selection, a good approximation for s is

$$s = ia/\sigma, \tag{2.9}$$

where i is the intensity of selection (standardized selection differential) and σ is the phenotypic standard deviation. Eqn. (2.9) is accurate as long as a is small relative σ (Falconer, 1981).

Stabilizing selection. The particular model of stabilizing selection to be analysed is the 'nor-optimal' model first used by Haldane (1954) and later by Latter (1960), Lande (1976) and Turelli (1984). With the optimum fixed at zero, the relative fitness is given by

$$W(X) = \exp(-X^2/2w^2), \tag{2.10}$$

where w is a measure of the strength of stabilizing selection. Increasing w implies weaker stabilizing selection. This equation is used directly in the Monte Carlo simulation to assign relative fitnesses to individuals. For single loci, Robertson's (1956) approximation for the change of gene frequency at one locus is used. This assumes that many loci contribute to the variance of the character whose mean is near equilibrium and can vary due to gene frequency changes at any of the loci contributing to the character. Robertson showed that a mutant allele behaves as under-dominant (i.e. the heterozygote is less fit than the homozygotes). The change of gene frequency at one locus under such stabilizing selection is approximately

$$\Delta q = a^2(q - 1/2)q(1 - q)/[4(w^2 + \sigma^2)], \tag{2.11}$$

where σ^2 is formally the phenotypic variance less the genetic variance contributed by the locus. The term $w^2 + \sigma^2$ is often called the strength of stabilizing selection and referred to as V_s. This is equivalent to a model of heterozygote inferiority in fitness where $s = a^2/[8(w^2 + \sigma^2)]$ is the fitness disadvantage of the heterozygote and there is a meta-stable equilibrium at $q = 1/2$. Importantly the strength of selection is proportional to the square of the allelic effect. Mutant genes are unconditionally deleterious and their

selection is similar to that of genic selection with $\Delta q = s^* q(1 - q)$, where

$$s^* = -a^2/[8(w^2 + \sigma^2)]. \tag{2.12}$$

2.3. Transition matrix method

With a suitable expression for the change of gene frequency from selection, Δq, it is possible to model selection of a mutant allele in a finite population using a transition matrix. This method allows the computation of the expected (and variance of) heterozygosity contributed by the mutant each generation in the population after it appears, and can also be used to compute the cumulative heterozygosity (and variance of heterozygosity) it contributes during its lifetime. At the core of the method is a transition matrix, **M**, which defines the transition probabilities of gene frequency. The use of such a matrix originates from work of Fisher and Wright and the process of change of gene frequency is often referred to as the 'Wright-Fisher stochastic process' (see Ewens, 1979, Ch. 3). The elements of the square matrix **M** are

$$m_{jk} = \binom{2N}{k}(q + \Delta q)^k\,(1 - q - \Delta q)^{2N-k} \qquad (0 \leq j,\ k \leq 2N), \tag{2.13}$$

where $q = j/(2N)$. Let $\mathbf{f}^T(t)$ denote the row vector with elements $f_j(t)$ which are the probabilities of the population having gene frequency $j/(2N)$ $(0 \leq j \leq 2N)$ at generation t. Thus for a new mutant $f_1(0) = 1$ and all other elements are zero. The vector $\mathbf{f}^T(t)$ at generation t $(t>0)$ is obtained from

$$\mathbf{f}^T(t) = \mathbf{f}^T(t\text{-}1)\mathbf{M}. \tag{2.14}$$

Let $\mathbf{h}$ denote a column vector whose elements are the expected heterozygosity at a locus with gene frequency $j/(2N)$ $(0 \leq j \leq 2N)$, so $h_j = 2(j/(2N))(1 - j/(2N))$. The expected cumulative heterozygosity, $H(a)$, contributed by a new mutant until it is fixed or lost is $H(a) = \sum_{t=0}^{\infty} \mathbf{f}^T(t)\mathbf{h}$. This can be computed as

$$H(a) = \mathbf{f}^{T*}(0)(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{h}^{*} \qquad (2.15)$$

where $\mathbf{I}$ is the unit matrix, $\mathbf{Q}$ is the square submatrix of $\mathbf{M}$ of dimension $2N-1$ defined by $q_{ij} = m_{ij}$, $(1 \leq i, j \leq 2N-1)$, and $\mathbf{f}^{T*}$ and $\mathbf{h}^{*}$ are the corresponding elements of $\mathbf{f}^{T}$ and $\mathbf{h}$ (Kemeny and Snell, 1960). A similar matrix equation can be derived for the variance of heterozygosity contributed by a mutant, but this was only used to check results of numerical analysis (Ch. 5).

The above describes one method for determining the equilibrium and dynamic behaviour of a mutant whose effect, a, appears in the transition matrix as an expression for change of gene frequency, Δq. In order to analyse the fates of mutations with effects sampled from a distribution occurring at all possible loci, numerical integration was used.

2.4. Numerical integration

With the assumption that the fates of each mutant are independent, selection responses, cumulative heterozygosity and cumulative variance of heterozygosity are all functions of mutant effect, a, integrated over $f(a)$. Numerical integration was used to compute these quantities using Simpson's rule. Consider the evaluation of the cumulative heterozygosity contributed by a mutant allele. The cumulative heterozygosity for a range of effects a in one size of population and one value of strength of selection was computed. The integration procedure was called with the scale of the procedure defined by ε and the shape defined by β. In order to reduce the amount of computation, the diffusion approximation was used to compute results for other population sizes, as the heterozygosity maintained is a function of Ns as long as N is not too small. Using $N = 80$, the approximation was found to be accurate, so results for other population sizes and a given strength of selection were computed by integrating using an appropriate value of ε. Convergence of the results of the integration procedure was checked by comparing two successive halvings of the a interval. The expected genic variance in the population for a distribution of mutant effects can be computed from,

$$V_g = 2N\lambda \int_{-\infty}^{\infty} (a^2/4)H(a)f(a)da. \qquad (2.16)$$

With highly leptokurtic forms (e.g. gamma with β = 1/4), problems occur in integration because much of the weight in the distribution is in the long tail, and the shape of the distribution changes very quickly near a = 0. Integration was therefore done over a number of ranges (typically 10) of a in order that a check could be made that sufficiently large values of a had been included, and most points (typically 257) were in the range nearest a = 0. Selection responses and variance of genetic variance were computed in a similar manner to (2.16). Details are given in Ch. 5.

2.5. Monte Carlo simulation

As the name implies, this type of simulation involves a computer program which generates random numbers to control the various processes operating in a simulated population. The random numbers are in fact pseudo-random and are generated by a special algorithm, a random number generator. This has one parameter, a seed integer, and generates a random real variable from a uniform distribution and a new seed. Thus, a computer program using this algorithm must, as part of its initialisation routine, read a seed. Each seed generates exactly the same sequence of random numbers, a useful feature for verification and debugging of programs. The number of possible seeds is sufficiently large and the characteristics of the random number generator are such that the use of different seeds makes runs effectively independent.

2.5.1. Principal parameters of the simulations

Before the simulation commenced, the values of a number of parameters were read: (1) Parental population size (N), and possibly progeny population size (T) (truncation selection only). (2) The number of pairs of chromosomes per individual. Usually, there was only one pair of chromosomes, but in some cases a large number of pairs of chromosomes without recombination was simulated which correspond to freely recombining loci with no intra-genic recombination and multiple alleles. (3) The length (L) of each chromosome in morgans. (4) The expected number of mutations per haploid chromosome per generation (λ). (5) The size of mutant effects expressed as $\varepsilon = [E(a^2)/V_E]^{1/2}$. The environmental variance, V_E, was scaled to unity. (6) The shape of the

distribution of effects of mutant alleles, defined by β. Most of the simulations are for β = 1/2. (7) The proportion of mutants of positive value (P). (8) The selection mode, either directional, stabilizing or neutral (no selection), and, in the case of stabilizing selection, the strength of selection (w). (9) The number of generations (t). (10) The number of replicates.

It will be noted that with the exception of (2) above (which is a rather special case results of which appear at only one place in the thesis), the concept of numbers of loci affecting the trait does not exist. Effectively, it was assumed that there was an infinite number of sites capable of affecting the trait and that these were evenly spread along the chromosome. The model is therefore similar to that described by Robertson (1970, 1977), except in that case, it was assumed the effects of segregating alleles were infinitesimally small, and there was no mutation.

2.5.2. Constitution of the population

The population consisted of an array of N parents and T progeny. With the exception of truncation selection the parental and progeny population sizes were the same. The progeny were generated from the parents each generation by a mating and recombination algorithm. The number of crossovers used to generate a recombinant chromosome was sampled from a Poisson distribution with parameter L. Each chromosome had an associated variable containing a count of the mutant sites. A chromosome was a two dimensional *real array*, each pair of elements representing a mutant which had two attributes, a position and a value. Generations were discrete, and mutations occurred before selection of parents to generate the progeny. New mutants had an immediate effect on the individual in which they occurred (they can be thought of as occurring in the gametes of the parents).

2.5.3. Generation of mutants

Each generation, individuals were 'mutated'. For each chromosome, the number of new mutants was sampled from a Poisson distribution with parameter λ. The effects of new mutants were sampled from a gamma distribution with shape parameter β and scale parameter ε. The sign of the mutation effect was assigned with probability P of being positive. The position of each mutant on the chromosome was sampled from a uniform distribution. The value and position of the mutant are assigned to the next element in the chromosome

array (mutants were not stored in any particular order), and the count of mutations was incremented. For practical purposes, each mutant occurred at a different site from those already segregating.

2.5.4. Selection

Two different selection procedures were used to generate the results depending on the selection mode. The results for directional selection either came from truncation selection or fertility based selection. For stabilizing selection, a fertility selection model was assumed.

Fertility selection. The parental and progeny population sizes were the same ($N = T$). With directional selection, the relative fitness of an individual was simply $W(X) = X + 1$. In this case the simulation of an environmental effect was unnecessary and the selection coefficient s was equivalent to the mutant effect, a. With stabilizing selection, the value of an individual was the sum of genotypic contributions plus a random environmental effect and the relative fitness is given by (2.10). Parents were selected for breeding with probability $w(X)/(\Sigma(W/N))$.

Truncation selection. The progeny population size was greater than that of the parents. The parents in the next generation were the N progeny of highest value and these were mated at random. The value of an individual was the sum of gene effects plus a random environmental contribution as above. Values and pointers to the array of progeny were sorted by the QUICKSORT algorithm (Knuth, 1968).

2.5.5. Measurement of effective population size

It is often useful to know the effective population size during the steady state. In order to measure N_e, each individual had a number (typically 6) of independently segregating pairs of neutral genes which did not contribute to the selection. Each generation, a random normal deviate of mean zero and variance V_M was added to the value of each gene. The steady state variance maintained on average at each locus (V_N) is related to the effective population size according to $V_N = 2N_e V_M$, so N_e was simply computed as $N_e = V_N/(2V_M)$.

2.5.6. Computation of results

Most of the time, the steady state behaviour of the population was investigated. The simulation started from an isogenic state and mutations were allowed to accumulate. In theory, by generation $6N$, the expectation of the genetic variance with neutral genes (the worst case) is more than 95% of its equilibrium value. At least this number of generations were ignored before the population was assumed to be at steady state. Thereafter, the following measurements were made at intervals of 10 generations: (1) Additive variances among parents and progeny. (2) Genic variance of progeny. (3) 'Neutral variance' of progeny to compute N_e. (4) Change of mean in the character to give response rate. (5) Skewness and kurtosis of genotypic values of progeny.

There was strong auto-correlation between successive measurements of means and variances, so the the standard error of the mean as computed from one run was biased downwards. In an earlier version of the software (Keightley and Hill, 1983), the means and standard errors were measured from one run only, but in the present version, the means and s.e.'s were estimated from independent replicates (i.e. start from isogenic state and ignore $>6N$ generations), so the estimate of the s.e.'s of the means were unbiased. In both methods, the estimates of the means were unbiased (except by failure to reach steady state, but this is negligible for long runs and in the presence of selection).

2.5.7. Fixation of mutants

Mutants which became fixed no longer contributed to the variance but used up computer time and memory, so to prevent the program slowing down, these were eliminated from the population by a 'garbage collection' routine. This was executed every eight generations (eight was found to be the optimal time between calls for most parameter sets). This routine sorted each chromosome in order of mutant values by QUICKSORT, recorded the mutants present on all chromosomes, then deleted them and recycled the free memory. Mutants which became lost were dealt with automatically by the progeny generation algorithm and no special routine was necessary to cope with these.

2.5.8. Random number generators

Uniform distribution. All the subsequently described random numbers rely on a good generator of uniform random numbers. In the simulation programs, the standard IMP80 generator (Stephens and Murison, 1982) was used which generates real random numbers from the uniform distribution in the range 0..1 ($U(0,1)$).

Poisson distribution. Firstly, a table, tab, of the cumulative density function of the Poisson distribution with parameter ξ was built with probability function

$$f(x) = e^{\xi}\xi^{x}/x! \quad \text{for } x = 0,\ 1,\ 2,\ \ldots$$

A random integer from the Poisson distribution was then generated by sampling a uniform, u from $U(0,1)$, then looking up the table and selecting the lowest value of j, such that tab(j) $\geq u$.

Normal distribution. The method described by Box and Muller (1958) was used to generate normal deviates of mean zero and s.d. x, $N(0,x)$. This directly transforms two independent $U(0,1)$ variates, $U1$ and $U2$ to a normal deviate, $N(0,1)$:

$$N(0,1) = [-2\ln(U1)]^{1/2}\cos(2\pi\, U2).$$

A second independent normal deviate from the same two uniforms can also be generated, but this was not implemented.

Gamma (1/2) distribution. This corresponds to a Chi-squared distribution with 1 d.f. (Kennedy and Gentle, 1980, Ch. 6). Random real numbers from the gamma distribution of scale parameter α and shape parameter 1/2 were generated by squaring a normal deviate $N(0,(1/2\alpha)^{1/2})$.

Coin toss. In order to randomly sample one or zero with equal probability, the individual bits of a random uniform constitute independent trials and these were used for this frequently called routine. This resulted in a worthwhile improvement in computation time compared to using one random uniform for

each coin toss.

Computer programs and their verification

A general practical problem with any computer programming is verification. With Monte Carlo simulation, the problem is particularly acute because the results are noisy and theoretical results are generally unknown for all parameter sets. Three approaches were used to verify the programs. (i) Structured development in a test-bed. The key routines were written in library modules, so a test-bed program could exercise the routines and verify their behaviour. (ii) Three independent programs were written which checked each other at the margins. These were for (a) no recombination; (b) variable recombination from zero to 20 morgans; and (c) free recombination between mutant sites. So, the variable recombination program with parameter $L = 0$ agreed with the results of the no recombination program. For large L, the results of the variable recombination program approached the results of the free recombination program. (iii) Running of limiting cases with known results from theory. For example, in the absence of selection all programs gave $V_g = 2NV_M$. Results could also be compared to those from the transition matrix for the case of $L \to \infty$.

The software was written in the IMP80 language (Stephens and Murison, 1982), a block structured procedural language developed by Edinburgh Regional Computing Centre (ERCC). IMP80 is similar to ALGOL and its derivatives and has many advantages over FORTRAN such as availability of recursion (used in QUICKSORT) and high level data structures. The programs were run on the ERCC's Amdahl V8 mainframe computer. Because of their multiple locus nature, the programs used very large amounts of computer time especially for large populations and high genomic mutation rates. In many cases the computing time per generation was proportional to $N^2\lambda$, so the computing time to reach a steady state was proportional to N^3.

Chapter 3

Directional Selection and Variation in Finite Populations

3.1. Introduction

In recent years there has been much interest in the production and maintenance of variation in populations by mutation, stimulated by the presence of abundant variation in natural and artificial populations at the protein and DNA levels. Also, the genome is now seen as a fluid entity with transposition a particularly potent force in generating molecular variability. Variation at the phenotypic level must also originate from mutation, but the rate at which such variation is generated has been thought to be slow. This belief was derived mainly from observations of experimental populations of Drosophila. For example the gain from new mutations in bristle score variation is of the order of one thousandth of the environmental variation per generation (discussed by Lande, 1976; Hill 1982b; see Ch. 1), and mutagenesis experiments have failed to produce large amounts of new variation in such quantitative traits (Clayton and Robertson, 1964; Kitagawa, 1967; Hollingdale and Barker, 1971).

Despite the apparent slowness of accumulation of new mutational variance, theoretical analyses of the interaction of mutation and natural selection in the absence of drift have shown that mutation may be a powerful force in maintaining variation in natural populations (Lande, 1976), although the extent predicted depends on assumptions in the model (Turelli, 1984). Theoretical studies in finite populations have concentrated on the combined effect of mutation and directional selection in influencing quantitative variability and selection response rates (Hill, 1982a,b). The equilibrium variance of a quantitative character is attained more quickly in the presence of selection than in its absence, and is highly dependent on population size. Thus in the early generations of a selection experiment or breeding programme the response from variation generated by new mutations is expected to be small. In later generations, however, the contribution to the total variation present and hence to the response can be very important, especially in large populations. The results of long-term selection experiments can be interpreted in light of these

analyses. The continued response after 120 generations of directional selection for increased pupal weight in Tribolium (Enfield, 1980), after at least 75 generations of selection for increased bristle score in Drosophila (Yoo, 1980), and after 76 generations for increased oil content in maize (Dudley, 1977) were likely to have been strongly influenced by variation arising while the experiment was proceeding.

Previously (Keightley and Hill, 1983) Monte Carlo simulation was used to investigate the effect of linkage on asymptotic selection responses from new mutations in small populations and it was concluded that the asymptotic response rate is little affected by linkage, especially for species with many chromosomes. Moreover, the variance of effects of mutants on the trait was not an important parameter of the model for asymptotic selection responses were little affected by whether the new mutational variance arose from a few genes of large effect or many genes of small effect. The effects of mutant alleles on the trait were, however, assumed to be symmetrically distributed about zero.

Here, a theoretical framework is developed to predict the amount of quantitative variation maintained and selection responses in finite populations from the simultaneous segregation of newly arising linked mutations. The analysis is based on the 'infinitesimal model' first used by Fisher (1918), which provides a reference point to allow a better understanding of the results of a more complex Monte Carlo simulation.

Cases where the effects of new mutations come from a skewed distribution are also investigated, i.e. mutations of increasing or decreasing effect on the character are not equally likely. Whilst it may be reasonable to assume that the effects of mutant are symmetrically distributed for Drosophila bristle number, it is likely that mutations affecting characters closely related to fitness are mainly detrimental (Mukai *et al.* 1972). The implications of the behaviour of the model are discussed in relation to theories of evolutionary advantages of recombination. In particular, conditions necessary for the operation of "Muller's Ratchet" (Muller, 1964), where the population mean can decline due to the fixation of recurrent mutations of negative effect are discussed.

3.2. Analysis of the infinitesimal model

3.2.1. Change of variance

The three quantities of interest are the additive and genic variances, and the disequilibrium covariance of the character. The additive variance is simply the variance of breeding values. The genic and disequilibrium components are given by (2.4) and (2.6) respectively. In an isogenic population, all three quantities V_A, V_g and ΣD are zero. With constant forces of mutation, selection and drift, they approach equilibria when the rate of loss of variation due to selection and drift is balanced by the rate of gain from new mutations. In a finite population the variances drift stochastically about the equilibrium. A prediction of the infinitesimal model is that the genic variance (V_g) is unaffected by selection (see Crow and Kimura, 1970, pp.236–239; Bulmer, 1980, Ch. 9) and therefore in an infinite population with recurrent mutation V_g never reaches an equilibrium. The equilibrium value of the additive genetic variance is affected by selection, mutation, drift and recombination. As a starting point, consider the effects of selection in an infinite population with free recombination in the absence of mutation.

Selection. The effect of selection on genetic variation with an infinitesimal model has been discussed by Bulmer (1971, 1980 pp. 153–154), and by Falconer (1981 pp.179–189). After one generation of selection of parents and breeding of progeny the total genetic variance in the population can be divided into fractions between and within full-sib families. With random mating, the between family component is one-half of the genetic variance among parents. Selection by truncation reduces the variance in the parents by a factor of $1 - h^2k$, where h^2 is the heritability ($h^2 = V_A/V_P$, the squared correlation between phenotype and genotype; with selection on an arbitrary index h^2 is replaced by the squared accuracy of the index), and k is a constant factor for the strength of selection. For truncation selection of a normally distributed population, $k = i(i - x)$, where x is the standardized deviation of the truncation point. Thus after one generation, the between family component of variance is $(1 - h^2k)V_A/2$.

Selection causes a reduction in the genetic variance between family means, which appears as a negative disequilibrium covariance component within

families. With free recombination the within family variance component is simply given by $V_g/2$ because recombination completely eliminates disequilibrium within families, but only half of the total genetic variance is initially present within full-sib families.

The total additive variance in the population after one generation of selection is obtained by adding the between and within family components,

$$V_{A,t+1} = (1 - h_t^2 k) V_{A,t}/2 + V_{g,t}/2. \qquad (3.1)$$

Recurrence relation (3.1) corresponds to equation (9.30) of Bulmer (1980). Its validity depends on a normal distribution of genotypic values in the progeny, since skew in the distribution can affect the amount of variation removed by selection, but the results of Bulmer (1980, Ch. 9) and Zeng (1987) indicate that, in many cases, skewness effects are small and can be ignored. Here, simulation is used to investigate possible effects of such skewness.

Finite population size and mutation. With a Poisson distribution of family size the expected reduction in the additive genetic variance in the population is by a proportion $1/2N_e$ in the absence of selection. With truncation selection, the within family variance is independent of the population size, but the expected reduction in the between family component is by a proportion $1/N_e$ due to sampling of parents with replacement. As with the genic variance, the additive variance increases by V_M units each generation from mutation. Equation (3.1) becomes

$$V_{A,t+1} = (1 - 1/N_e)(1 - h_t^2 k) V_{A,t}/2 + V_{g,t}/2 + V_M. \qquad (3.2)$$

Linkage. This does not affect the variance between family means after one generation of selection, but affects the within family variance by reducing the amount of variation recovered from the disequilibrium covariance component due to recombination between loci. If c is the recombination fraction between loci i and j, the disequilibrium remaining in the within family component is given by

$$2\sum_{i\neq j}\sum D_{ij,t}(1 - 2c_{ij})a_i a_j = \sum D_t - 2\sum(D_t c), \tag{3.3}$$

where ΣD_t is defined by (2.6) and similarly $2\Sigma(D_t c) = \sum_{i\neq j}\sum D_{ij} a_i a_j c_{ij}$. More generally (3.2) becomes

$$V_{A,t+1} = (1 - 1/N_e)(1 - h_t^2 k) V_{A,t}/2 + [V_{g,t} + \sum D_t - 2\sum(D_t c)]/2 + V_M. \tag{3.4}$$

The recurrence relation for the disequilibrium component is obtained by combining equations (2.7) and (3.4) to give

$$\sum D_{t+1} = \sum D_t - \sum D_t/(2N_e) - \sum(D_t c) - h_t^2 k(1 - 1/N_e)(V_{g,t} - \sum D_t)/2. \tag{3.5}$$

3.2.2. Asymptotic variance

As $t \to \infty$ for finite N, the variances reach expected equilibrium values about which they fluctuate stochastically due to sampling. For free recombination (3.2) can be re-expressed (ignoring second order terms) as a quadratic in V_A,

$$V_A^2(1 + k) + V_A(V_E - 2N_e V_M) - 2N_e V_M V_E = 0. \tag{3.6}$$

Thus, V_A is a function of mutation rate and population size as their product NV_M.

For complete linkage ($c = 0$ for all possible pairs of loci) a quadratic in V_A is obtained by combining (2.7) and (3.4),

$$V_A^2 N_e k + V_A(V_E - 2N_e V_M) - 2N_e V_M V_E = 0. \tag{3.7}$$

Again second order terms are ignored. Here, the relationship of V_A to V_M and

N is not as simple as in (3.6), but is a function of $N_e V_M$ and $N_e k$.

When recombination is finite, the simultaneous recurrence relations (3.4), (3.5) and (2.5) do not appear to have a simple solution. Their properties were investigated by iterating until steady state was achieved with initial values of V_g, V_A and ΣD set at zero, as would be the case in an isogenic population. The effect of a finite chromosome length was modelled by dividing the chromosome into a large number of equivalent segments (typically 100) and calculating the recombination fraction and hence the disequilibrium contribution from each possible pair of segments. This method exactly models the case of infinitesimally small effects as the number of segments approaches infinity, but increasing the number of segments beyond 100 made almost no difference to the results. The total amount of recombination was specified by L, the map length of the chromosome, and Haldane's (1919) mapping function was used to related recombination fraction to map distance (l) between pairs of loci: $c = [1 - \exp(-2l)]/2$. Previous analyses (Avery and Hill, 1979) indicate that other models relating recombination fraction to map length (e.g. with crossover interference) make little difference to the behaviour of this type of model.

3.3. Results

3.3.1. Comparison of the simulation and the infinitesimal model

Predictions of the expected value of V_A from the Monte Carlo simulation and from the infinitesimal model for varying population size are shown in Fig. 3.1. The value of V_M was $10^{-3} V_E$ and in the simulation a range of values of size of gene effects, ϵ, was compared with corresponding values λ to satisfy (2.1).

With free recombination, the infinitesimal and simulation models are in good agreement. The agreement is close even with relatively large selective values and few mutants (ϵ = 0.4). The disequilibrium present in the populations simulated can be estimated by subtracting the observed V_A from the genic variance (approximately $2NV_M$ in the infinitesimal case). With free recombination the amount of disequilibrium is small.

With complete linkage, the curves for different values differ substantially, larger values of ϵ giving higher predictions of V_A. The infinitesimal model is a poor

predictor for complete linkage especially when effects are large, but it also overestimates V_A when $\varepsilon \to 0$. The over-estimation can be explained by the presence of negative skew in the distribution of genotypic values of individuals (Table 3.1). Negative skew leads to a greater loss of variance each generation than predicted by the constant factor $i(i - x)$, and hence a lower expected V_A.

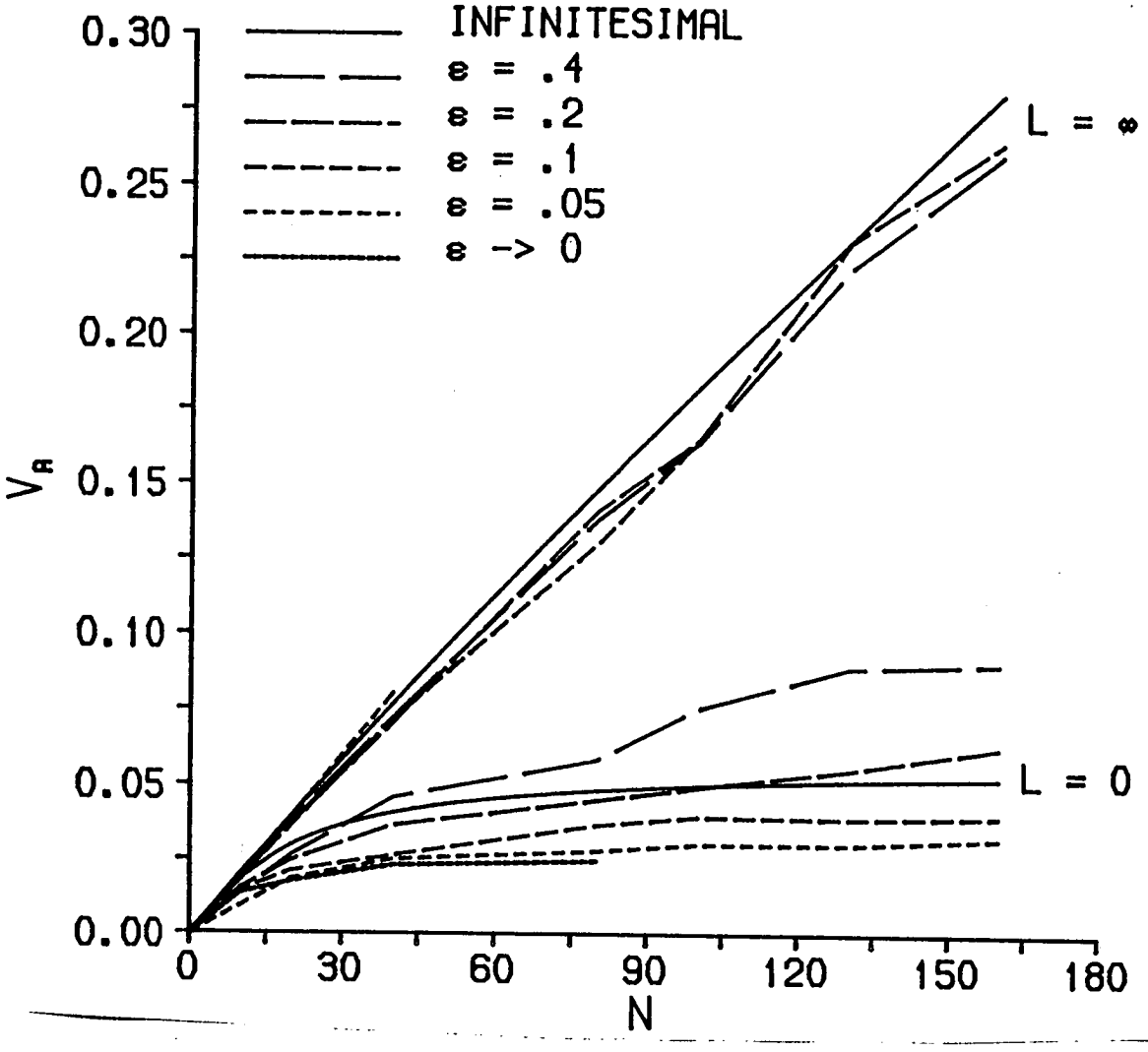


Fig. 3.1 The equilibrium variance is shown for the infinitesimal and simulation models for various population sizes and 50% truncation selection. $V_M/V_E = 10^{-3}$. In the simulation model, a range of values of ε and corresponding λ were used, with mutations coming from a symmetrical 'reflected gamma' distribution ($P = 0.5$).

Table 3.1.

The equilibrium skewness of genotypic progeny values, computed as $g_1 = [\Sigma(G\text{-MEAN})^3/N]/V_A$.

	Population size (N)			
	10	20	40	80
ϵ	Equilibrium skewness (g_1) among progeny			
→0	-0.0794	-0.147	-0.180	-0.184
0.05	-0.0592	-0.105	-0.149	-0.145
0.1	-0.0533	-0.0540	-0.0717	-0.0381
0.2	0.0142	0.0220	0.0465	0.0852
0.4	0.0289	0.0568	0.138	0.0735

Results are given for the case of no recombination ($L = 0$). $V_M/V_E = 10^{-3}$ and mutants come from a symmetrical reflected gamma distribution. Fifty percent truncation selection was simulated.

The effects of a finite amount of recombination are shown in Fig. 3.2. The simulation and infinitesimal models agree at the free recombination limit but there is an increasing discrepancy at low recombination fractions. At the population sizes simulated, most of the effect of linkage is eliminated by one or two crossovers per chromosome per generation. The results are in agreement with those of Keightley and Hill (1983) which used a fertility model of selection rather than the present viability model.

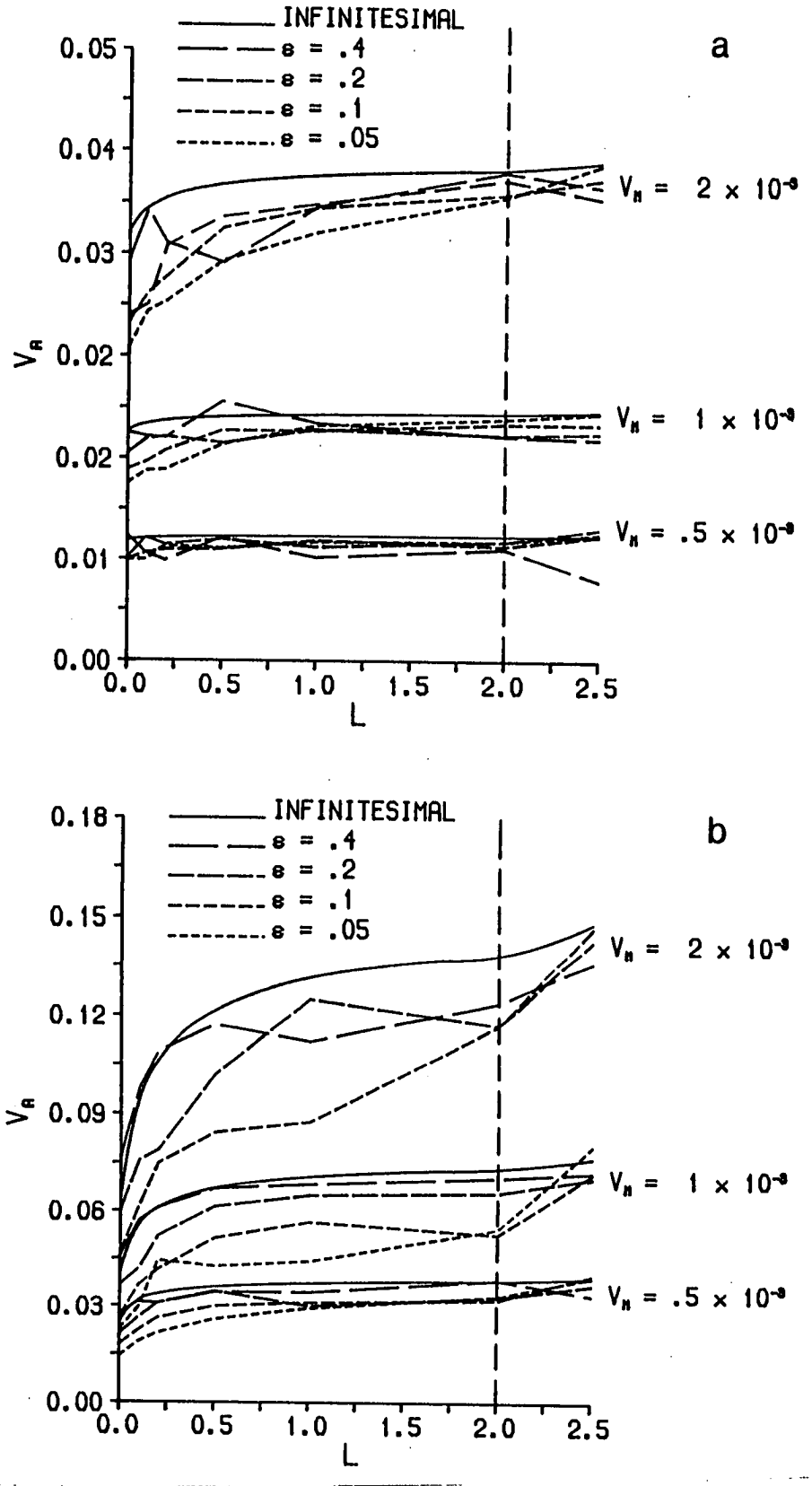


Fig. 3.2. Equilibrium variance (V_A) predicted from the simulation and infinitesimal models for three different values of mutational variance are plotted for different chromosome lengths (L) with 50% truncation selection. (A) N = 10; (B) N = 40.

3.3.2. Asymmetrical distribution of mutational effects

Predictions of V_A from simulations of different population sizes using the 'reflected gamma' distribution are plotted in Fig. 3.3 for free recombination and values of P representing cases where mutants are mostly negative (P = 0.1), positive (P = 0.9) or symmetrically distributed (P = 0.5). The results show that the expectation of V_A is higher than the infinitesimal prediction (also shown in the figure) when mutants have predominantly positive effects and lower when most are negative. As the expected value of mutational effects approaches zero, however, the results approach the infinitesimal prediction. In the limit all the effects become infinitely small and the models coincide. With finite effects of mutants, there are two reasons for the discrepancy from the infinitesimal prediction. Firstly, most mutants of negative effect are lost almost immediately and contribute little to the variance maintained (Hill, 1982b), while many mutants of positive effect are fixed and contribute substantially to variance especially when at intermediate frequencies. The variance maintained is proportional to the fraction of the mutational variance contributed by positive effects, $E^+(a^2)$ (Hill, 1982a). Secondly, the mutation pressure generates skew in the distribution of genotypes (negative or positive depending on the sign of the mutations) since the density function of mutational effects is itself skewed (see Table 3.2).

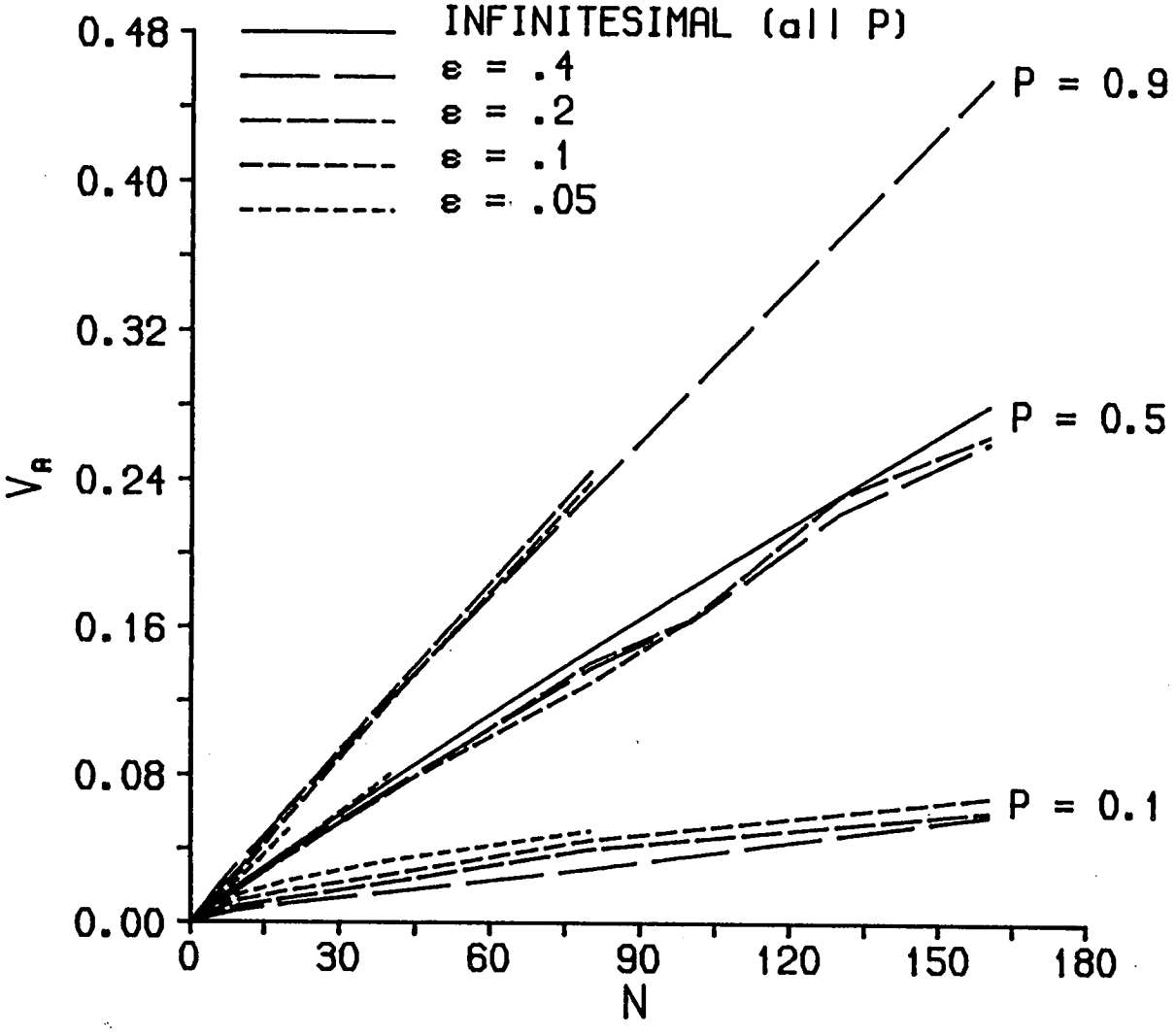


Fig. 3.3. Equilibrium variance (V_A) in the simulation and the infinitesimal models for various population sizes and three values of P (proportion of mutants positive). Free recombination, otherwise parameters as in Fig. 3.1.

Table 3.2.

The equilibrium skewness of progeny genotypic values computed as in Table 3.1.

	Proportion of mutants positive (P)		
	0.1	0.5	0.9
ε	Equilibrium skewness (g_1) among progeny		
0.05	-0.118	0.0312	0.0785
0.1	-0.285	-0.0121	0.0539
0.2	-0.699	0.0514	0.134
0.4	-1.67	0.0773	0.385

Results are given for a reflected gamma distribution of mutants, $V_M/V_E = 10^{-3}$ and free recombination. The population size (N) = 20 and 50% truncation selection was simulated.

3.3.3. Response to selection

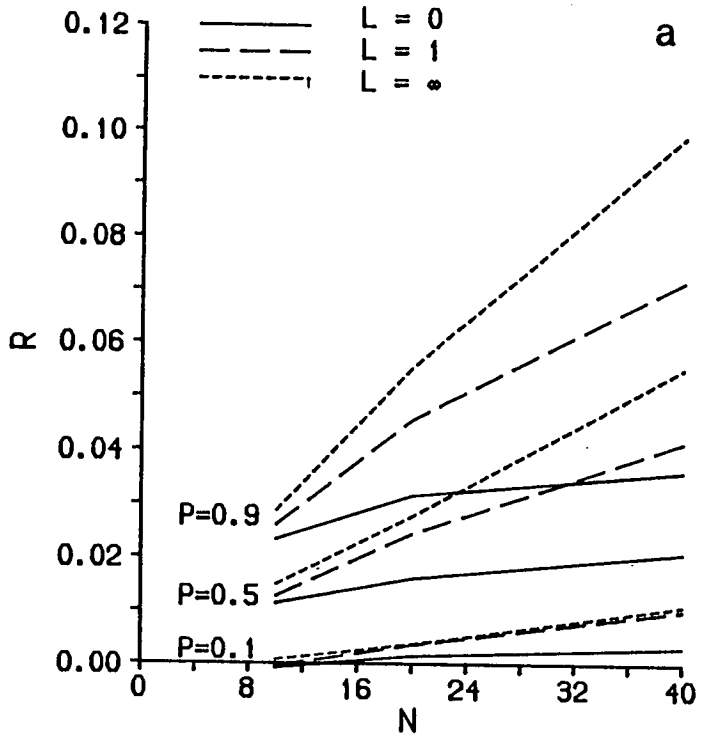
The response is given by $R = iV_A/\sigma$: if the distribution of genotypes and environmental effects are normal and independent of one another, the regression of A on P is linear (e.g. Falconer, 1981). When the mean value of mutant effects is non-zero, there is an additional change in mean, $\Delta_m = 2\lambda E(a)$, due to the mutational pressure. In the gamma distribution reflected about zero with shape parameter $\beta = 1/2$, $E(a) = \varepsilon(2P - 1)/\sqrt{3}$ so

$$\Delta_m = 2\lambda\varepsilon(2P - 1)/\sqrt{3}. \tag{3.8}$$

The responses to selection for various population sizes with both free and zero recombination and three values of P are plotted in Fig. 3.4. Since Δ_m is independent of population size but V_A is highly dependent, net responses become negative in small populations if most mutations are deleterious ($P = 0.1$).

Restating (2.1) and (3.8), $V_M \propto \lambda\varepsilon^2$ and $\Delta_m \propto \lambda\varepsilon$. It is clear that for a given V_M, as the expected magnitude of effects decreases ($\varepsilon \to 0$) and hence the number of mutations increases, Δ_m must increase. Thus, if the new mutational variance is due to a large number of small negative effects, the mean value of a population will decline faster than if the mutational variance is due to a small number of large effects.

Paradoxically, the effects of linkage, i.e. the difference between response rates for free recombination and complete linkage, are most severe when most mutations are positive ($P = 0.9$ in the figure). Both positive and negative mutants interfere with each other's fixation probabilities (Hill and Robertson, 1966), but when most mutations are of positive value, there are more segregating so linkage effects are more important.



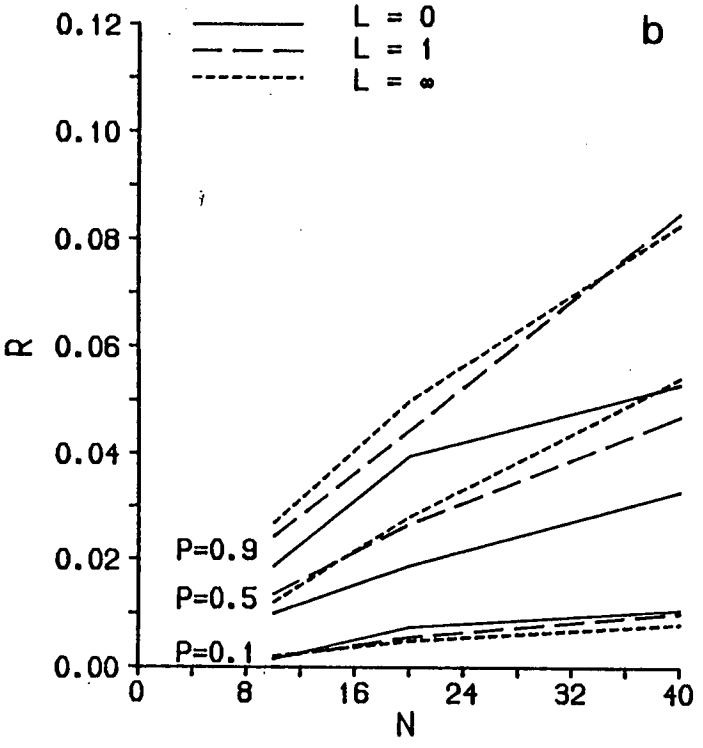


Fig. 3.4. Response rates where mutant effects have a reflected gamma (1/2) distribution with three values of P (proportion positive) in various population sizes. $V_M/V_E = 10^{-3}$ and 50% truncation selection. (A) $\epsilon = 0.1$; (B) $\epsilon = 0.4$.

3.4. Discussion

With V_M/V_E of the order of 10^{-3}, it is clear that with directional selection and free recombination, disequilibrium is a minor part of the total variation present, and a small number of crossovers is a good approximation of free recombination. Using the 'infinitesimal model' an analytical solution has been developed for the equilibrium genetic variance under the joint effects of mutation, linkage and selection in a finite population. The model agree well with the Monte Carlo simulations both where effects are small ($\varepsilon \to 0$), and also where mutational effects are relatively large. This behaviour is consistent with the results of Hill (1982a) where independent genes give V_A of $2N_eV_M$, irrespective of the mutational distribution. The higher fixation probability of mutants of large effect and their higher contribution to the variance in the character is nearly exactly balanced by their shorter fixation times and fewer number when compared to genes of small effect.

When mutants come from an asymmetrical distribution the behaviour is not as simple. As $\varepsilon \to 0$ for any population size, the equilibrium additive variance is essentially the same as predicted by the infinitesimal model. This is true irrespective of the selection regime, selection only generating disequilibrium. With larger effects, V_A becomes dependent on the proportion of mutants of positive effect (P). When the value of most mutational effects exceeds σ/N_ei, V_A is approximated by $4PN_eV_M$. This is so because the fixation probability of such mutants (and therefore the probability that they will reach intermediate frequencies and contribute substantially to the population variance) is proportional to a and independent of N_e. The number of mutants appearing in the population is, however, proportional to N_e. For small effects or in small populations, terms for which $|a| < \sigma/N_ei$ become more important and V_A approaches the infinitesimal prediction of $2N_eV_M$. Fig. 3.3 shows, however, that mutant effects must become very small or N_e very small before V_A is much different from $4PN_eV_M$.

With an asymmetrical mutational distribution, the distribution of genotypic values becomes skewed in the same direction as the skew of the mutational distribution. The skewness of the genotypic distribution affects the amount of variation removed by directional selection, so equilibrium variances are affected by such skewness. The simulations show that skewness is more important

when mutant effects are large. It is notable that directional selection also generates skewness (positive) in the genotypic distribution (Bulmer, 1980 Ch. 9).

The behaviour of the model with linked genes is also strongly affected by skewness in the genotypic distribution. As $\varepsilon \to 0$ in the simulation, the infinitesimal model overestimates the equilibrium genetic variance. The likely explanation is a negatively skewed mutational distribution generated from the loss of all but the best haplotype and the presence of a 'tail' of individuals of lower value from mutation. This tendency to generate negative skewness is partially opposed by truncation selection generating positive skewness as mentioned earlier. The effect of linkage in generating skewness has been noted in earlier two locus studies (Hill and Robertson, 1966). Where effects are large, and therefore for a given V_M fewer mutations occur per generation, genes behave more as if they were independent and therefore higher V_A is maintained.

As a consequence of a negatively skewed mutational distribution, the rate of fixation of deleterious genes can exceed the rate for beneficial mutants and the population mean can decline; an effect corresponding to "Muller's Ratchet" (Muller, 1964). A number of conditions have been identified as necessary for the operation of the ratchet: (1) small population size since the fixation of deleterious mutants depends on chance; (2) many mutants of small effect (as opposed to a few of larger effect) since the 'mutation pressure' on the population mean is greater in this case; (3) tight linkage since less standing variation is available to oppose the mutation pressure. Linkage is also more important with many small effects (cf. Fig. 3.1).

Somewhat surprisingly the simulations show that linkage has most influence where most mutants are of positive value, and linkage effects can all but disappear when most are negative (cf. Fig. 3.4). The explanation, however, is simple: deleterious genes almost never become fixed, while positive mutants become fixed with probability, proportional (if independent) to α. In this latter case, however, linked positive mutants present simultaneously in the population can form unfavourable repulsion combinations leading to a reduction in fixation probabilities.

Chapter 4

Mutation-Stabilizing Selection Balance in Finite Populations

4.1. Introduction

Many quantitative characters show considerable heritable variation in natural populations. Explaining how such genotypic variation is maintained has been one of the most important and controversial problems of population genetics. The problem arises because of the widespread belief that stabilizing selection in which the fittest individuals have values of the trait near some optimum is ubiquitous in nature, but selection for an intermediate optimum is expected theoretically to deplete genetic variability (Robertson, 1956) and has been shown to do so experimentally (Gibson and Bradley, 1974; Kaufman, Enfield and Comstock, 1977). There is a certain irresistibility in arguments for the presence of intermediate optima: for example the date of egg laying in many Northern passerine birds apparently has an optimum dependent on the availability of caterpillars for the young which are only present for a brief period in early summer (Lack, 1968). Some of the most compelling evidence for selection for intermediate optima in natural populations comes from comparisons between sibling species of Drosophila where parallel latitudinal clines for various traits have been shown to exist (David and Bocquet, 1975; Hyytia *et al.*, 1985). The observation of an intermediate optimum at any single trait considered alone is, however, not in itself evidence of stabilizing selection as Robertson (1973) and others (Falconer, 1981; Rose, 1982; Hill and Keightley, 1988) have emphasized, because negative correlations between characters under directional selection can also generate such optima.

Mutations are the basic source of all heritable variation, but can a balance between mutation and selection alone explain the maintenance of observed high levels of genetic variation? This is an important question because variation in quantitative traits is believed to be the 'raw material' of evolution. Such variation also provides the basis for responses to artificial selection, and it is important in understanding the allelic effects and gene frequencies contributing to the variation being utilised.

The work of Clayton and Robertson (1955) suggested that mutation is a weak

force in generating quantitative variation. Recent work on directional selection and mutation has, however, indicated that long-term selection responses may in part be due to mutations occurring after commencement of the experiment (Hill, 1982b). Lande (1976) focused on mutation-stabilizing selection balance and, by fitting experimentally estimated parameters to a specific model, concluded that high levels of genetic variation can be maintained in the presence of strong stabilizing selection. The assumptions of Lande's "continuum of alleles" model were based on results of Kimura (1965), i.e. new mutants have effects that differ only slightly from those pre-existing with the result that the distribution of allelic effects segregating at a locus is approximately normal. Although Lande included an analysis of the effect of linkage, the formulae obtained were essentially the same as those of Kimura (1965).

In a more recent review of the experimental data, Turelli (1984) questioned the appropriateness of the Kimura-Lande (KL) model since with experimentally measured mutation rates there are unlikely to be more than two alleles segregating at the loci affecting the trait. Also, the effects of new mutations are likely to be larger than the existing range of variation at the loci with the consequence that the distribution of allelic effects segregating at the locus is non-normal, an *assumption* critical to Lande's model. With an assumption of lower per locus mutation rates, Turelli (1984) obtained a formula for the equilibrium variance in the population which contrasts markedly with the KL result, i.e. the equilibrium genetic variance is independent of the effects of mutants on the trait, but depends only on the total number occurring per generation. This result of Turelli with a "House of Cards" approximation (Kingman, 1978) was obtained earlier for a two allele model by Latter (1960) and Bulmer (1972), and a similar answer has been subsequently obtained for a five allele model by Slatkin (1987). Burger (1986, see also Burger, Wagner and Stettinger, 1988) has generalized the analysis of the KL continuum of alleles model, and shown that Turelli's "House of Cards" result is a very good approximation for the KL model over a very wide range of parameters.

In this chapter, the mutation-stabilizing selection problem is analysed for finite populations. Finite population size is likely to be important because, although population sizes in nature can be very large, they are seldom constant and an equilibrium model of the maintenance of genetic variation must consider past fluctuations in effective population number. Also, more importantly, the results

of Robertson (1956) showed that the strength of stabilizing selection on an allele is proportional to the square of its effect on the character. Previously, there have been two major analyses of this problem, by Latter (1970) and by Bulmer (1972). Latter's analysis was a finite population extension of Kimura's (1965) continuum of alleles model and assumed a normal distribution of allelic effects segregating at the loci affecting the trait. Bulmer's model assumed up to two alleles segregating and equal forward and backward mutation rates at each locus, with the effects of substitution the same at each locus. A formula was obtained for the equilibrium variance with parameters: effective population number, mutant effect on the trait, strength of stabilizing selection and mutation rate.

Here, the influence of the shape of the distribution of the effects of new mutant alleles on the variation maintained in the character is investigated. The effects of new mutants on quantitative traits vary because they can occur at different places within genes (e.g. flanking sequences, introns, intron-exon boundaries, 'silent' third positions, promoters, active sites, other coding regions), but also because they can occur at genes whose functions vary within the biochemical and developmental system.

Most of the analysis concentrates on a model of two alleles per locus. This is similar to Bulmer's (1972) analysis. The consequences of allowing for the possibility of the presence of more than two alleles is investigated using Monte Carlo simulation. The effects of linkage on the equilibrium genotypic distribution are investigated using Monte Carlo simulation.

4.2. Results

4.2.1. Approximate analysis

Insight into the behaviour of a new mutant allele can be gained from extending Latter's (1960) two allele treatment. The variance contributed by a mutant allele is given by $V_g = a^2 q(1 - q)/2$. Assuming the mutant has a sufficiently large deleterious effect (strictly large Ns) that the mutant is eliminated before reaching appreciable frequency, the variance contributed is approximated by $V_g \simeq a^2 q/2$. The expected variance after selection is therefore $V_g' \simeq a^2(q + \Delta q)/2$, and substituting Robertson's (1956) expression (2.11) for the change in gene frequency, this is approximated by

$$V_g' \simeq a^2 q[1 - a^2/(8(w^2 + \sigma^2))]/2. \qquad (4.1)$$

Thus, the expected change in genetic variance from one generation of stabilizing selection is by a factor $1 - a^2/[8(w^2 + \sigma^2)]$. The expected proportional change in variance from drift is by a factor $1 - 1/2N_e$. Considering mutants of equal effect occurring at rate λ per haploid genome per generation, the expected increment in the variance of the character from mutation is, from (2.1) $\lambda a^2/2$, so a recurrence relation describing the balance between selection, drift and mutation may be written down:

$$V_{g,t+1} = V_{g,t}[1 - a^2/(8(w^2 + \sigma^2))][1 - 1/2N_e] + \lambda a^2/2. \qquad (4.2)$$

Ignoring second order terms, this gives a solution at equilibrium ($t \to \infty$) of

$$V_g = \lambda N_e a^2/[1 + N_e a^2/(4(w^2 + \sigma^2))]. \qquad (4.3)$$

The same equation has been obtained independently by Burger, Wagner and Stettinger (1988) using a different, heuristic, argument. In terms of the proportion of variance, $2N_e V_M = N\lambda_e a^2$, maintained with no selection, the variance is

$$V_g = 2N_e V_M/[1 + N_e a^2/(4(w^2 + \sigma^2))]. \qquad (4.4)$$

As N becomes infinite (4.3) reduces to

$$V_g = 4\lambda(w^2 + \sigma^2), \qquad (4.5)$$

the well known result of Latter (1960), Bulmer (1972) and Turelli (1984). Where mutant effects on the trait come from a distribution, the equilibrium variance can be computed by integrating (4.3) over a density function of mutant effects,

but result (4.5) still holds. Equation (4.3) also shows that the equilibrium variance is (to first order) a function of $N_e a^2/(w^2 + \sigma^2)$.

4.2.2. Two allele model

Heterozygosity as a function of $N_e s$. The expected cumulative heterozygosity contributed by a mutant during its lifetime as a function of $N_e a^2/(w^2 + \sigma^2) \propto N_e s$ is illustrated in Fig. 4.1, computed using the transition matrix (eqn. (2.15)). $H(a)$ is bounded by the upper value of two where drift dominates and the lower value of zero where selection causes immediate elimination of the new mutant. The results in the following sections which give examples of V_g for different types of gamma distributions are all functions of the result in Fig. 4.1 and were generated by integrating numerically over this function with weighting according to the distribution of mutant effects. In retrospect, these results could have been computed by evaluating Bulmer's (1972) formula for the genetic variance in a finite population with mutation and stabilizing selection (see Ch. 5). The results from the transition matrix, however, were in close agreement.

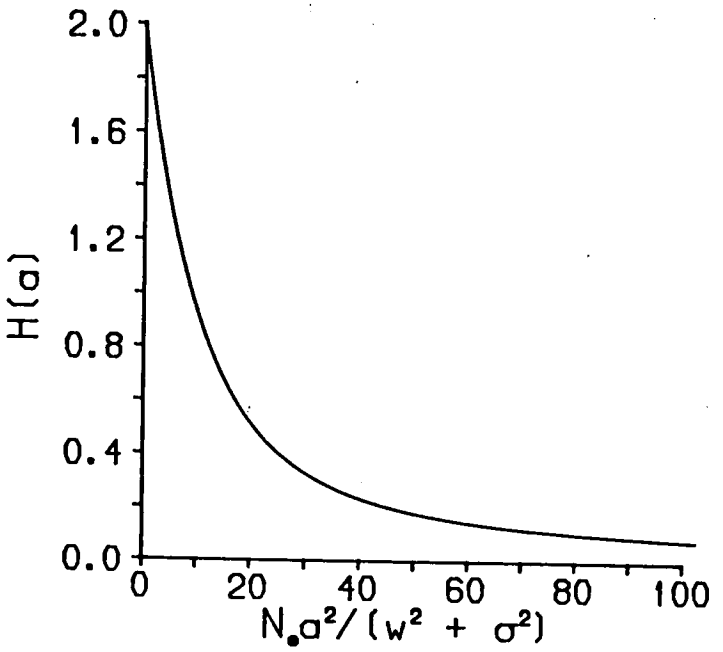


Fig. 4.1. Heterozygosity maintained during the lifetime of a mutant as a function of $N_e a^2/(w^2 + \sigma^2)$, derived using the transition matrix.

Variance maintained with genes of equal effect. Fig. 4.2 shows the variance maintained as a proportion of that predicted in an infinite population (cf. equation (4.5)) as a function of $N_e a^2/(w^2 + \sigma^2)$. The graph compares the results from the transition matrix and evaluation of (4.3). With increasing effects of drift ($N_e s \to 0$), the variance maintained approaches zero; and as the effects of selection become more important (increasing N_e or $a^2/(w^2 + \sigma^2)$), the relative variance maintained approaches the asymptote of one. Interestingly, the results from the transition matrix indicate a maximum greater than the infinite population variance. If the effects of drift and selection are not too strong, the frequency of some mutants can approach the meta-stable point ($q = 0.5$) where the expected change in gene frequency is zero. This possibility is not accounted for by equation (4.3) which assumes that the selection coefficient is constant and at its maximum. The presence of the maximum in Fig. 4.2 was confirmed by Monte Carlo simulation with equal mutant effects in a multi-locus model (results not given).

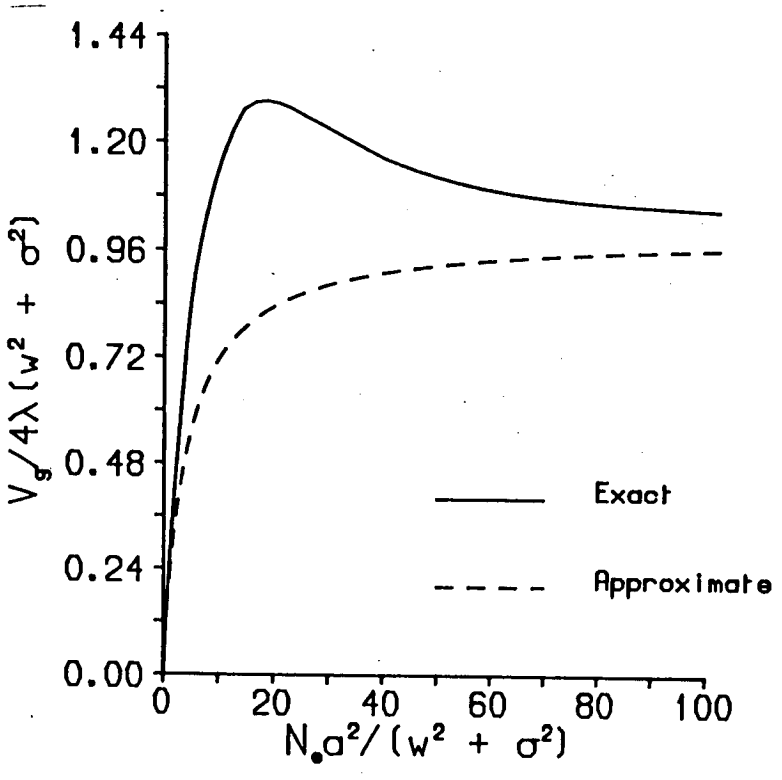


Fig. 4.2. Predicted equilibrium genetic variance, V_g, expressed as a proportion of that predicted for an infinite population, $4\lambda(w^2 + \sigma^2)$, plotted against $N_e a^2/(w^2 + \sigma^2)$. The curves compare predictions from the transition matrix (exact, ——) and from evaluation of eqn. (4.3) (approximate, ----).

The influence of the scale of the mutational distribution. This is illustrated for varying population size in Fig. 4.3 for the case of a gamma distribution of effects with shape parameter $\beta = 1/2$. The curves were generated by evaluating cumulative heterozygosities using the transition matrix and numerically integrating (2.16). The different curves relate V_g for a fixed value of V_M with various mutation rates (λ) and with corresponding values of sizes of effects (ε) to satisfy (2.1). With $\lambda \to \infty$ and infinitesimally small mutant effects the expected V_g is simply $2N_e V_M$ as obtained by Clayton and Robertson (1955). This is the upper bound of the maintained variance. The initial trajectory of all the other curves is also $2N_e V_M$, but each slowly approaches the asymptotic value given by (4.5) as N_e increases.

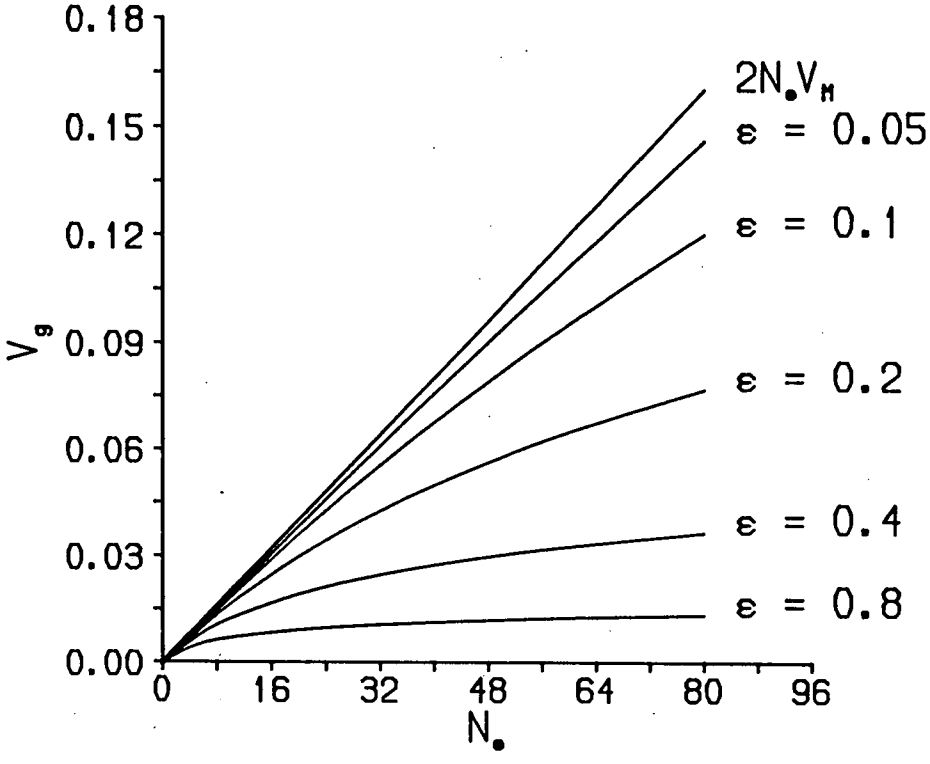


Fig. 4.3. Predicted equilibrium genetic variance maintained plotted against population size N_e. The curves were generated from the transition matrix, assumed a value of $V_M/V_E = 10^{-3}$, a gamma distribution of mutant effects, with $\beta = 1/2$ and compare results for a range of $\varepsilon = [E(\alpha^2)/V_E]^{1/2}$ and corresponding mutation rate λ. The strength of stabilizing selection is given by $w^2 = \sigma^2 = 1$.

The influence of the shape of the mutational distribution. Fig. 4.4 shows the variance maintained in a finite population as a proportion of that which would be maintained in an infinite population, expressed as a function of $N_e\varepsilon^2 V_E/(w^2 + \sigma^2)$. The curves relate different gamma distributions of mutant effects ranging from equal ($\beta \to \infty$) to a highly leptokurtic form ($\beta = 0.25$) (see Fig. 2.1). The result for equal effects is also shown in Fig. 4.2. Clearly, the shape of the distribution has a strong influence on V_g, and with highly leptokurtic forms, the approach to the asymptote is exceedingly slow. Curves for other values of shape parameter also have maxima. A normal distribution of mutant effects would correspond to a gamma distribution reflected about zero with shape parameter $\beta = 1.75$ (Hill and Rasbash, 1986), so on Fig. 4.4 its curve would lie between $\beta = 1$ and $\beta = 2$.

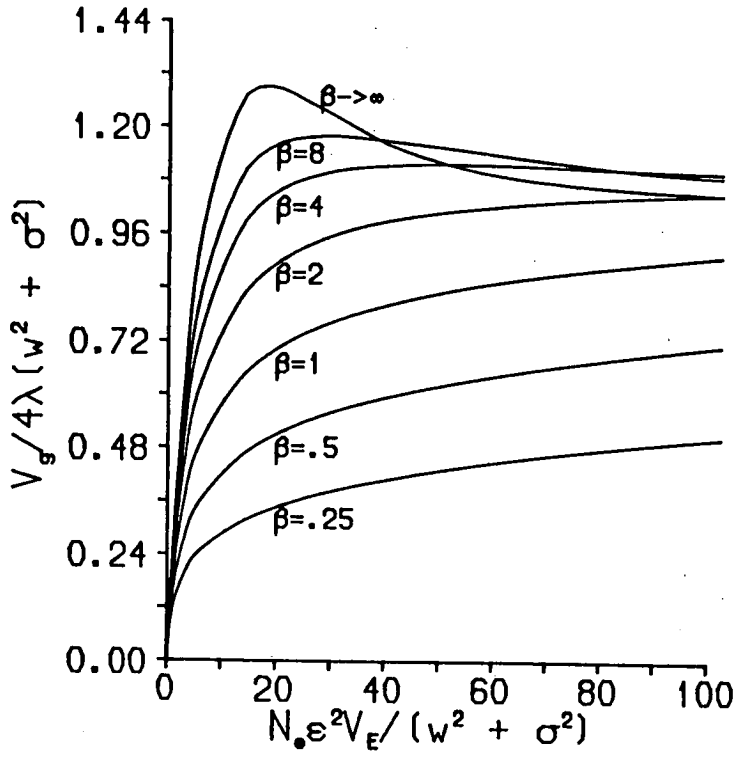


Fig. 4.4. Predicted equilibrium genetic variance, V_g, expressed as a proportion of that predicted for an infinite population, $4\lambda(w^2 + \sigma^2)$, plotted against $N_e\varepsilon^2 V_E/(w^2 + \sigma^2)$. The curves were generated by integrating the gamma function for a range of shape parameter, β, from equal effects ($\beta \to \infty$) to an extremely leptokurtic distribution ($\beta = 0.25$). Results from the transition matrix.

Linkage. Using the transition matrix method, it is only feasible to work out the expectation of the variance contributed by a single locus. In order to incorporate linkage, Monte Carlo simulation was used. Figure 4.5 compares the additive genetic variation maintained from mutations occurring on chromosomes of three different lengths in populations of varying sizes. With free recombination, the simulation results are in agreement with the predictions from the transition matrix obtainable from Fig. 4.4, which confirms that in the absence of linkage, mutants can be treated independently to approximate the behaviour of the system. Linkage leads to a reduction in maintained additive variance which is greatest with many mutants of small effect ($\varepsilon \to 0$) and its effects are virtually absent with the larger values. As in the case of directional selection (Ch. 3), a small amount of recombination eliminates most of the effects of linkage, but is less effective as N increases.

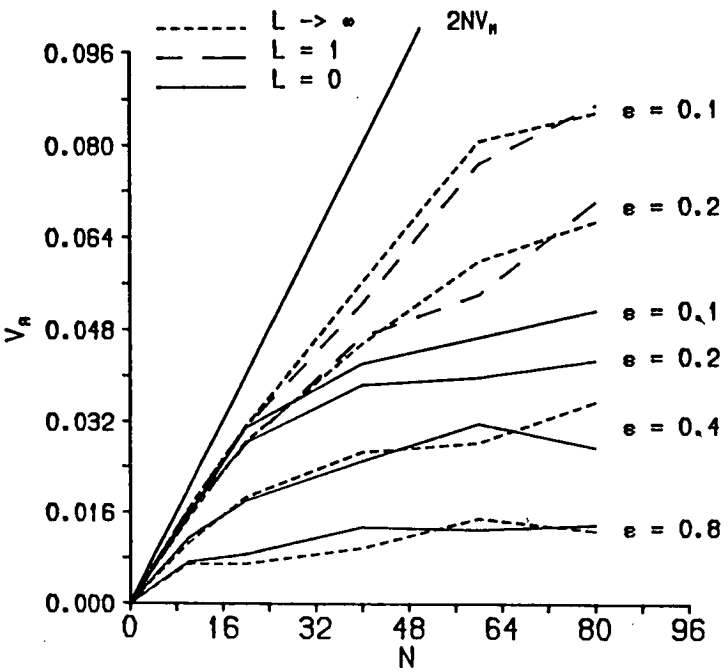


Fig. 4.5. Equilibrium additive variance plotted against population size, N, for a gamma distribution of mutant effects with shape parameter $\beta = 1/2$ and a range of values of $\varepsilon = [E(a^2)/V_E]^{1/2}$ and corresponding values of λ to give $V_M/V_E = 10^{-3}$. The curves are from Monte Carlo simulation and compare three lengths of chromosome, L, where L is the number of map units. The strength of selection, is given by $w^2 = \sigma^2 = 1$. The standard errors of the simulation points are less than 10% of their mean.

4.2.3. Multi-allele model

The above analysis is a finite population treatment which assumes that only two alleles can segregate at any locus. In this respect it is similar to the models of Latter (1960) and Bulmer (1972) and to Turelli's (1984) "House of Cards" approximation (henceforth referred to as the LBT models), with the additional assumption that mutant effects are randomly sampled. The models of Kimura (1965) and Lande (1976) (KL Gaussian models) assume that mutant effects differ only slightly from those already segregating with the consequence that the steady-state distribution of allelic effects at the locus is approximately normal. As Turelli (1984) pointed out, there is a fundamental discrepancy between the behaviour of the two types of model. Using Kimura's analysis, the equilibrium variance in an infinite population at a locus (V_{gL}) is

$$V_{gL} = [2V_{ML}(w^2 + \sigma^2)]^{1/2}, \tag{4.6}$$

where V_{ML} is the mutational variance input at the locus. This result can be derived by a different route. Assume at steady state a large number of alleles generates a normal distribution of allelic effects segregating at a locus. The variance maintained at the locus in a finite population can be obtained from the recurrence

$$V_{gL,t+1} = V_{gL,t}(1 - 1/2N_e)(1 - V_{gL,t}k/2\sigma^2) + V_{ML}, \tag{4.7}$$

because the variance at the locus is reduced each generation by the factor $(1 - 1/2N_e)$ by drift and $(1 - V_{gL}k/2)$ by selection, where k depends on the strength of selection and is the proportion of the phenotypic variance in the unselected individuals (Bulmer, 1980; Falconer, 1981 p.180). With stabilizing selection and a normal distribution of phenotypic values, $k = w^2/(w^2 + \sigma^2)$. For an infinite population, (4.7) reduces to Kimura's (1965) formula (4.6) (ignoring second order terms). Equation (4.7) also gives a solution for a finite population which is a quadratic in V_{gL}.

$$V_{gL}^2 N_e k + \sigma^2 V_{gL} - 2N_e V_{ML}\sigma^2 = 0. \qquad (4.8)$$

This formula is similar to that obtained using the same assumptions by Latter (1970) and is similar to equation (3.7) for the equilibrium variance under directional selection contributed by a chromosome in the absence of recombination.

Fig. 4.6 compares the equilibrium genetic variance maintained for a range of population sizes for three different models using a gamma distribution of mutant effects with shape parameter $\beta = 1/2$: (1) Gaussian: the variance was computed from the solution to (4.8). This corresponds to the KL prediction. (2) 'Two allele': the variance was computed from the transition matrix and numerical integration, for a model of two alleles only per locus. (3) 'Multi-allele': the variance was computed by simulation of n discrete freely recombining loci with no intra-genic recombination, so the number of alleles which can segregate at any locus is not limited. Also shown is the variance maintained by neutral genes which is simply $2NV_M$. The main points to note from this Figure are: (a) All three models agree at small population sizes and small mutant effects where drift is dominating. (b) The simulation of multiple alleles agrees with the KL Gaussian prediction only when mutant effects are small ($\varepsilon = 0.1$) and the mutation rate per locus is high ($\mu = 2X10^{-3}$). (c) Otherwise, with decreasing mutation rate, but correspondingly increased magnitudes of mutant effects (e.g. $\varepsilon = 1.6$ and $\mu = 7.81x10^{-6}$), the simulation agrees better with the two allele model. The simulation illustrates the difference between the KL approximation and Turelli's (1984) "House of Cards" approximation. With the number of loci chosen for this example ($n = 100$), the KL prediction hardly differs from neutrality. A larger number of loci would allow for a smaller standard deviation of the distribution of mutant effects for mutation rates per locus in line with those experimentally measured (Mukai and Cockerham, 1977). In this case, all three models would agree more closely at the population sizes shown, but as population sizes became much larger, would diverge as in Fig. 4.6 as the effect of selection becomes stronger relative to drift.

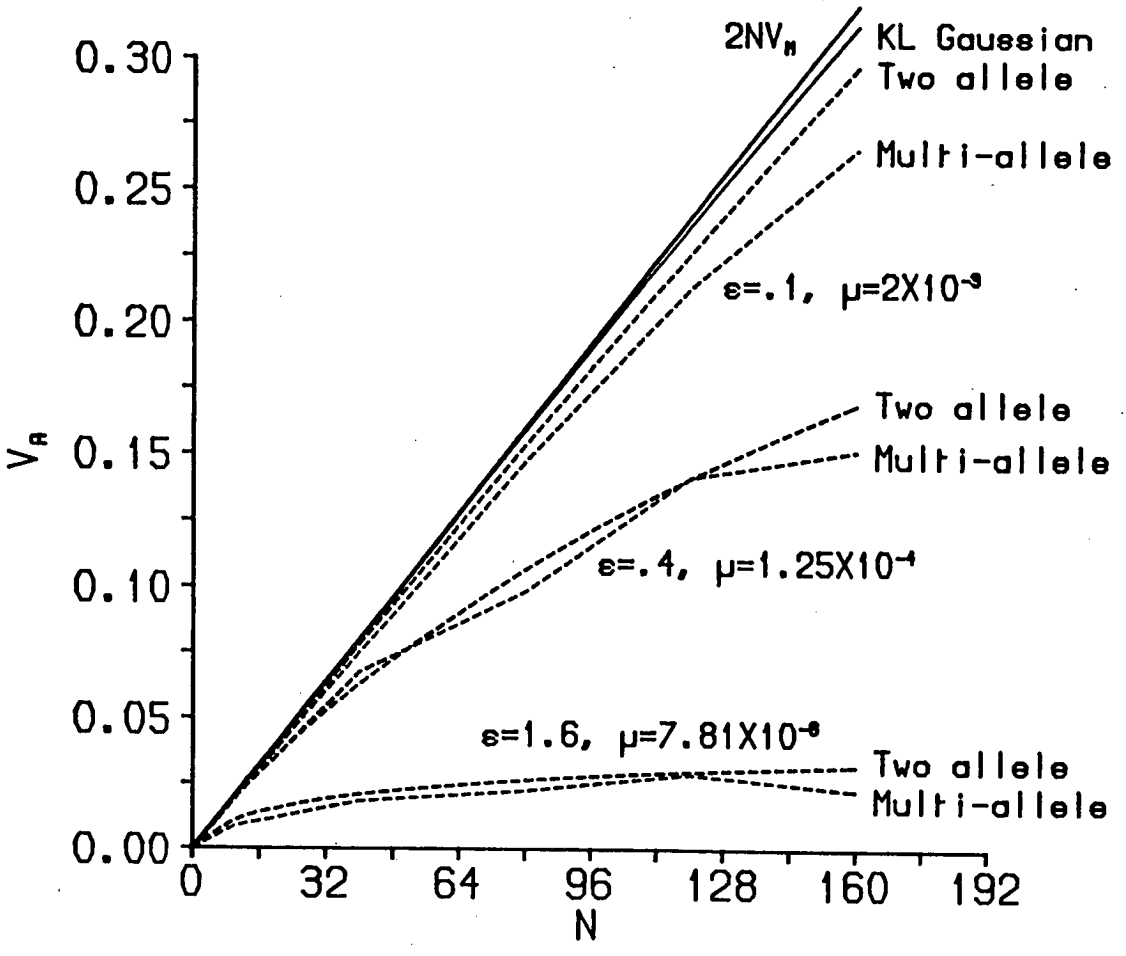


Fig. 4.6. Equilibrium additive variance plotted against population size, N, for three different models (see text). Also shown is the variance predicted for no selection, $2NV_M$. The simulation used 100 equally mutable freely recombining loci with mutation rates, μ, as shown, corresponding values of ε so that $V_M = n\mu\varepsilon^2/2 = 10^{-3} V_E$ and mutant effects sampled from a gamma distribution with $\beta = 1/2$. The strength of stabilizing selection is given by $w^2 = 18\sigma^2 = 19$. The standard errors of the results from the simulation are less than 5% of their mean.

4.3. Discussion

Stabilizing selection and drift. Most attention has been given to a model where the mutation rate is sufficiently low or the population size sufficiently small that two alleles segregate at each locus. The consequences of such a model with infinite population size have been investigated previously (Latter, 1960; Bulmer, 1972) and an important conclusion was that the equilibrium genetic variance, V_g, is essentially independent of the effects of mutants on the trait, but depends only on the number of new mutants per generation. As a consequence of its independence on the effects of mutants, in an infinite population V_g is independent of the shape of the mutational distribution. In a finite population, V_g is also proportional to the mutation rate, but in contrast is also highly dependent on the magnitudes of the effects of mutants. The variance maintained is a function of population size, the effect on the character and strength of stabilizing selection according to $N_e a^2/(w^2 + \sigma^2)$. The variance contributed by a mutant during its lifetime is at a maximum when the combination of parameters $N_e a^2/(w^2 + \sigma^2) \simeq 25$, and is about 30% greater than that in an infinite population, namely $4(w^2 + \sigma^2)$. So, for example, a mutant of effect $a = 0.1$ under stabilizing selection of strength $w^2 + \sigma^2 = 20$ would contribute during its lifetime about 30% more variance in a population of $5X10^4$ than in an infinite population. This effect occurs because near neutral mutants are able to drift to intermediate frequencies where the strength of selection is weakest, but mutants of larger effect tend to be eliminated almost immediately by selection. The maximum is a consequence of the multi-locus nature of the system. If no other genes were segregating when a mutation occurred, its fate would depend on the relation of the optimum phenotype to the population mean.

Distribution of mutant effects. In a finite population V_g is highly dependent on the shape of the distribution of mutant effects. Using gamma distributions a wide range of possible mutant distributions has been modelled ranging from all effects equal to a highly leptokurtic form where most mutants are of tiny effect, but most of the mutational variance, V_M, is contributed by a few genes of large effect. With such a distribution, in contrast to when effects are equal, V_g increases very slowly with decreasing effect of drift. This slow approach to the asymptote is best understood by considering a fixed mutational distribution and selection, but increasing population size. The mutants of large effect which

contribute most to V_M do not contribute substantially to V_g since they are quickly eliminated by selection, but the many mutants of small effect eventually contribute substantially to V_g with increasing population size because they remain as nearly selectively neutral until N_e becomes very large.

An implicit assumption of the analysis has been a symmetrical distribution of mutant effects. If there are few mutants segregating, or mutant effects are small, asymmetry does not influence the variance maintained, and this was confirmed by simulation (results not shown). The simulations also showed that in general, slightly less variance is maintained with a skewed distribution of mutant effects, because the population mean is moved away from the optimum and selection is thereby stronger against most new mutants. Also, the genotypic distribution becomes skewed (Ch. 3) and selection tends to remove more variation than it would from a symmetrical distribution.

Linkage. With stabilizing selection the extreme genotypes are less favoured, so there is a tendency for repulsion genotypes to persist and coupling genotypes to be eliminated. The result is that the additive variance is less than the genic variance because there is a 'hidden' negative disequilibrium component. The simulations show that disequilibrium increases with increasing mutation rate and with population size because there are more mutants segregating. Also, for the same reason, linkage has more influence when most of the mutational variance is contributed by many genes of small effect than a few genes of large effect.

The results show that recombination is very efficient at eliminating such influence of disequilibrium. In the examples shown, one crossover is sufficient to give results almost indistinguishable from free recombination. These simulations were, it should be emphasized, extreme cases with relatively strong selection and all the mutants appearing on one chromosome. The results from the simulation are relevant to the appropriateness of the two allele model (next section), because the simulation is an infinite sites model with no distinction between alleles and loci. The efficiency of even a small amount of recombination in eliminating linkage disequilibrium (cf. Ch. 3) implies that the two allele model is a good enough approximation because mutants occurring close together on the chromosome can be regarded as being either at the same or at different loci.

Appropriateness of the Model. Most of the analysis has been restricted to segregation of only two alleles per locus. A model has also been investigated which resembles more closely those of Kimura (1965) and Lande (1976) in which the effects of new alleles are assumed to be small relative to the existing variance at the locus at which many alleles segregate, so the asymptotic distribution of allelic effects at a locus is approximately normal. Turelli (1984) has questioned the appropriateness of the KL model since the mutation rate per locus is unlikely to sustain sufficient standing variation at a locus for the assumptions in the model to be valid. Simulation results (Fig. 4.6) with finite populations support Turelli's objections. The analytical result of Kimura (equal to first order to Lande's) is only in agreement where the mutation rate per locus is exceptionally high and mutant effects small. With mutation rates per locus closer to experimentally obtained estimates i.e. 10^{-4}–10^{-5} (Mukai and Cockerham, 1977; Turelli, 1984), the two allele model provides a good approximation.

4.3.1. Implications

The maintenance of genetic variation is a central problem in population biology, and the question of whether a mutation–stabilizing selection balance can maintain the observed levels of heritable variation has been frequently addressed (e.g. Lande, 1976, 1980; Turelli, 1984, 1985). The results show that with finite population size, the shape of the mutational distribution has a strong influence on the genetic variance maintained under mutation–selection balance. There is little information concerning the shape of the distribution of mutational effects for any character, but can an informed guess be made from insights into biochemistry and molecular biology? In principle all mutants, no matter where they occur in the genome, must have at least some effect on all characters, albeit very small. The interactive nature of metabolism, where the fluxes and metabolite pool concentrations are systemic properties dependent to a greater or lesser extent on all enzymes in the 'metabolic map' (Kacser and Burns, 1973, 1981) tells us that there must be hundreds, if not thousands, of enzymes, variation in the activities of which will affect any character which is in some way controlled by the metabolism of the organism. Evidence for functional constraint in the genome (Kimura, 1983 Ch. 7) at such sites as introns, silent (non-replacement) sites within coding sequences, and gene flanking sequences, suggests that there are many places in the genome

capable of producing some small phenotypic effect. Thus, it can be argued *a priori* that the distribution of mutant effects on complex quantitative characters is highly leptokurtic: most mutants are either of such trivial effect or so 'distant' from the character that they have almost no effect at all, but there is a smaller class of genes, more directly capable of influencing the trait with mutants of relatively large effect. The total number of mutants affecting a character is therefore high, much higher than an experiment designed to count polygenes would detect, but the effect of most of them is very small (see Robertson, 1967).

The difficulties in estimating the number and effects of mutants influencing a quantitative character are highlighted by the following illustrative example. Assume by genetic means only mutants showing an effect on the character of at least one-half of a standard deviation can be detected and the standard deviation of the mutational distribution is 1.6 units. If all effects were equal, then the genetic test would detect all the new mutants. If, however, the mutational distribution were more extreme, for example gamma with shape parameter $\beta = 1/4$ (see Fig. 2.1), then only 21% of the new mutants would be detected but they would contribute most of the variance (96%).

With this consideration in mind, estimates of the number of new mutants per generation affecting various quantitative traits in maize (Sprague, Russell and Penny, 1960; Russell, Sprague and Penny, 1963) seem rather high. These experiments gave estimates for *detectable* mutants per generation of about 0.06 implying, with, say, a mutation rate per locus of 10^{-5}, many thousands of loci at which mutations give sufficiently large effects to be detected. Such experiments, however, might now have to take into consideration the possibility of induction of "mutator" genes (McClintock, 1950) in these crosses caused by movement of transposable elements known to be capable of affecting quantitative traits (Mackay, 1987). The rates of mutation may vary widely between populations as results of T.F.C. Mackay (personal communication) suggest, due to varying transposition rates.

Predicting Maintained Heritability - Assigning Values to Parameters. The important parameters are the mutational variance input per generation, the shape and scale of the mutational distribution, the strength and mode of operation of natural selection, and effective population size. As implied earlier, information is scarce on values of most of these parameters relevant to natural

populations. If, however, it is assumed that: $V_M/V_E = 10^{-3}$; the character is affected by a fairly extreme distribution of effects (i.e. gamma with $\beta = 1/4$); most variance is contributed by mutants of fairly large effect (e.g. $\epsilon = 0.4$ and therefore the mutation rate per genome $\lambda = 0.0125$ implying c.400 loci each mutating at c.$0.3X10^{-5}$); and a "typical" value of $w^2 = 20\sigma^2$ (Turelli, 1984), then with $N_e = 10^3$, the maintained heritability would be about 21%; with $N_e = 10^4$, the maintained heritability would be about 33%, but there would be less than half of the genetic variance that would be present in an infinite population. Thus, on the face of it, mutation-stabilizing selection balance is an attractive candidate for explaining the observed levels of heritable variation over a wide range of effective population sizes. Such calculations become less attractive when we consider the problems in estimating the strength of natural selection and in justifying the single character model of stabilizing selection. Such aspects have been discussed in detail by Turelli (1984, 1985).

Mode of Action of Natural Selection. The effect of pleiotropy is to reduce the genetic variation maintained since, for example, the selection coefficient against a mutant if stabilizing selection acts independently on each character, i, is proportional to $\Sigma[a_i^2/(w_i^2 + \sigma_i^2)]$. Clearly, the analysis could easily be extended to include pleiotropy and Fig. 4.1 still applies with the horizontal axis labelled as $N_e\Sigma[a_i^2/(w_i^2 + \sigma_i^2)]$ reflecting the selection acting on all the other characters. If mutant effects on each trait are uncorrelated and there are k traits, then the variance maintained for each is reduced in proportion to $1/k$ (Turelli, 1985).

Pleiotropic gene action is likely to affect other characters subject to stabilizing selection, or to affect characters such as fertility and viability more closely connected with fitness *per se*. In the former case, the selection experienced by a mutant allele will be proportional to the square of the effect on the pleiotropically related character, but in the latter case the pleiotropic selection is directional and is proportional to the allelic effect to first order (see Hill and Keightley, 1988). Thus, especially when allelic effects are small, the selection due to the effect on fitness dominates and even small negative correlations with fitness-related characters are likely to have a large impact on the heritability maintained.

It seems therefore that in order to fully understand the maintenance of variation in a quantitative character the bivariate distribution of mutant effects on that character and on fitness is a critical parameter. Since the present

analysis suggests that mutants of small effect are likely to be more important in maintaining variation, the more accessible part of the mutational distribution may be of less interest. We are some way from a satisfactory understanding of the mechanisms of maintenance of variation in polygenic traits.

Chapter 5

Further Analysis of The Stabilizing Selection Model

5.1. Introduction

Many selection experiments use as their source material samples from caged insect populations (often Drosophila) which have been in the cage environment for many generations. Samples are taken from the cage and separate selection lines started. The response patterns of individual lines and the variation in response among lines are obtained. Similarly, selection of artificial populations is often with species which have been under domestication for many generations and there are often independent selection lines. Such experiments should provide information on the underlying genetic basis of quantitative variation.

Many characters in natural (and perhaps in artificial) populations are thought to have intermediate optima. A popular model of selection with an intermediate optimum is stabilizing selection. This has intuitive appeal, there is some evidence for its operation in nature and the model is amenable to analysis. Previous analyses of the consequences of stabilizing selection in natural populations have been concerned with the genetic variance maintained when the character is at or near the optimum in which case it is necessary to invoke mutation to maintain genetic variation. Two kinds of model have been analysed. The first was proposed by Crow and Kimura (1964) and involves loci controlling variation in a quantitative trait at which mutations appear, the values of which differ only slightly from the previous allelic state. Kimura (1965) analysed this model, derived a formula for the equilibrium genetic variance of a locus at a mutation-stabilizing selection equilibrium, and showed that the equilibrium distribution of allelic effects segregating at the locus is normal. The model was further analysed by Lande (1976) who argued that it predicts that substantial variation can be maintained even with strong stabilizing selection.

The second type of model differs from the first because the effects of mutant alleles can be large. Turelli (1984) analysed a "House of Cards" model, which was originally proposed by Kingman (1978). The critical assumption which differed from Kimura's (1965) analysis above was that the effect of a mutant

allele swamps existing variation at a locus controlling genetic variation in the trait. This gives different qualitative predictions of the variance maintained at the locus at equilibrium and agrees with two allele analyses of Latter (1960) and Bulmer (1972). Turelli argued that a "House of Cards" model of the mutation process is appropriate for natural situations.

With the exception of Bulmer (1972) the analyses described above have been of infinite populations. This has an important consequence for the equilibrium probability distribution of allele frequencies at a locus influencing the trait. With stabilizing selection in populations near equilibrium, mutations are unconditionally deleterious (Robertson, 1956). In an infinite population the equilibrium probability distribution of allele frequencies is therefore highly leptokurtic, i.e. mutant alleles are almost always very rare, and intermediate allele frequencies are absent. This affects the consequences of a shift in the optimum. Barton and Turelli (1987) analysed the dynamics of the population mean and variance after a change of the optimum and showed an accelerating rise in the mean, slowing down as it approached the new optimum, and a rise in the genetic variance because some previously deleterious mutant alleles became advantageous and were selected to intermediate frequencies where they contributed more substantially to the genetic variance. In some cases, the variance fell again close to its original value (i.e. alleles became fixed) and in others a new equilibrium was reached with a higher variance. The existence of such multiple equilibria was predicted by Barton (1986). A change of optimum is similar to imposing directional selection on a character previously subject to stabilizing selection.

In a finite population, alleles are able to drift in frequency, so the equilibrium probability distribution of allele frequencies becomes less leptokurtic than described above. Bulmer (1972) derived an expression for the probability distribution of allele frequency at a mutation-stabilizing selection-drift balance for the case of up to two alleles per locus. Here, expressions are derived for the variance of the genetic variance among independent lines at a mutation-selection-drift balance. The results of previous investigations of the problem using neutral models (Bulmer, 1976; Avery and Hill, 1977; Lynch and Hill, 1986) are compared. The consequences of a change from stabilizing selection to directional selection on the genetic variance of a character in a finite population are investigated. These results together with patterns of response to directional selection from Monte Carlo simulation of populations

previously under stabilizing selection are compared to response patterns obtained from experiments published in the literature.

5.2. Definitions

The finite population with recurrent mutation and the methods used to analyse it are describe in Ch. 2. The model is used to investigate the effect on the genetic variance of a change from stabilizing selection to directional selection. There are therefore two phases of selection, a stabilizing selection phase followed by a directional selection phase. From diffusion theory, in the first phase the steady state probability distribution of allele frequency with recurrent mutation is a function of $N_1 s_1$, where N_1 is the effective population number in the first phase and s_1 is the selective value of the mutant allele. With stabilizing selection s_1 is frequency dependent (Robertson, 1956) and is approximated by $s_1 = (q-1/2)a^2/[4(w^2 + \sigma^2)]$ or $s_1^* = -a^2/[8(w^2 + \sigma^2)]$ when the mutant allele is rare. The steady state distribution of allele frequency with recurrent mutation and stabilizing selection is therefore a function of the parameter $N_1 E(s_1^*) = -N_1 \varepsilon^2 V_E/[8(w^2 + \sigma^2)]$. In the second phase of directional selection, the pattern of change of gene frequency is the same on a scale of t/N_2 when the parameter $N_2 s_2$ is constant, where N_2 is the effective population size in this phase and s_2 is the selection coefficient of the favourable allele. With truncation selection the selection coefficient is approximately $s_2 = ia/\sigma$, where i is the intensity of selection, so the directional selection phase is parameterised by $N_2 E(s_2) = N_2 iE(|a|)/\sigma$.

5.3. Results

Before analysing the consequences of directional selection on a character previously subject to stabilizing selection, consider the variance of the genetic variance in a 'base population' which is assumed to be under stabilizing selection.

5.3.1. Single Locus Analysis

Using diffusion theory (Kimura, 1969), Bulmer (1972) derived the density function of gene frequency, q_i, at a locus under stabilizing selection with equal forward and backward mutation rates, μ, between two possible alleles

$$f(q_i) \propto \exp[(-4\Phi q_i(1 - q_i)][q_i(1 - q_i)]^{\Theta-1}. \quad (5.1)$$

where $\Theta = 4N_e\mu$, $\Phi = N_e a^2/[8(w^2 + \sigma^2)]$, and N_e is the effective population size. It follows that the expected heterozygosity at the locus is

$$E[q_i(1 - q_i)] = [\int_0^1 \exp(-4\Phi q_i(1 - q_i))(q_i(1 - q_i))^{\Theta} dq]/ \quad (5.2)$$
$$[\int_0^1 \exp(-4\Phi q_i(1 - q_i))(q_i(1 - q_i))^{\Theta-1} dq].$$

This was shown to reduce to

$$E[q_i(1 - q_i)] = I(4\Phi, \Theta+1)/I(4\Phi, \Theta), \quad (5.3)$$

where $I(x, y)$ is a function of the complete beta function, $B()$, and the confluent hypergeometric function, $M()$

$$I(x,y) = B(y, 1/2)(1/4)^{y-1/2} e^{-1/4x} M(1/2, y+1/2, 1/4x). \quad (5.4)$$

Assuming Θ is small (i.e. ignoring back-mutation) and integrating over the distribution of gene effects, $f(a)$, by expanding the confluent hypergeometric function as a series, it can be shown that the expected genic variance is

$$E(V_g) = N\lambda \int [(\sum_{i=0}^{\infty} \Phi^i/((2i+1)i!))/e^{\Phi}] a^2 f(a) da, \quad (5.5)$$

(Abramowitz and Stegun, 1965 Ch. 13). Bulmer's (1972) analysis can be extended to derive a formula, with similar assumptions, for the variance of the genic variance at a locus with recurrent mutation among independent lines. It follows from (5.1), (5.2) and (5.3) that

$$E[q_i^2(1 - q_i^2)] = I(4\Phi, \Theta+2)/I(4\Phi, \Theta). \qquad (5.6)$$

Assuming $\Theta \to 0$, by similar analysis the variance of the genic variance among lines is

$$V(V_g) = (N_e\lambda/12)\int[(\sum_{i=0}^{\infty}3\Phi^i/(i!(4(i+1)^2-1)))/\ e^{\Phi}]a^4 f(a)da. \qquad (5.7)$$

Eqns. (5.5) and (5.7) can be evaluated easily by iteration on a computer and converge readily. They were checked against results obtained from a transition matrix and were found to agree almost exactly. Eqn. (5.7) also agrees with results from the Monte Carlo simulation (Table 5.1).

Table 5.1

Comparison of predictions of variance of genic variance derived from (5.7) (diffusion theory of independent genes) and Monte Carlo simulation.

N	Theory $V(V_g)\times10^4$	Simulation $V(V_g)\pm1$s.e.$\times10^4$
10	1.94	1.95±0.08
15	2.92	2.85±0.17
20	3.89	3.68±0.13
30	5.83	5.65±0.18

The parameters of the simulation were $\lambda = 0.2$, $\varepsilon = 0.1$, a reflected gamma distribution of mutant effects with shape parameter $\beta = 1/2$, and no selection.

Two limiting cases are of particular interest.

Neutrality, $\Phi \to 0$. From (5.5) the expected genic variance is $E(V_g) = N_e\lambda E(a^2) = 2N_e V_M$. The variance of the genic variance from (5.7) is

$$V(V_g) = N_e \lambda E(a^4)/12, \tag{5.8}$$

in agreement with Lynch and Hill (1986) who used a different derivation. The coefficient of variation of V_g is therefore

$$CV(V_g) = [(E(a^4)/E^2(a^2))/(12N_e\lambda)]^{1/2}. \tag{5.9}$$

Strong selection, $\Phi \to \infty$. From (5.5) the expected genic variance is $E(V_g) = 4\lambda(w^2 + \sigma^2)$ (see Latter, (1960); Bulmer, (1972); Turelli, (1984); Ch. 4). The variance of the genic variance among lines from (5.7) is

$$V(V_g) = 4\lambda(w^2 + \sigma^2)^2/N_e. \tag{5.10}$$

Both the expectation and the variance of the genic variance with strong stabilizing selection (large $N_e s$) are therefore independent of the magnitude of the effects of mutant alleles. The coefficient of variation of V_g is

$$CV(V_g) = (1/4N_e\lambda)^{1/2}. \tag{5.11}$$

Comparison of (5.9) and (5.11) shows that for new mutant alleles of equal effect, the coefficient of variation varies by only a factor of $1/\sqrt{3}$ between cases of weak and strong selection. The shape of the distribution of effects of mutant alleles becomes important as selection becomes weak. Fig. 1 shows the fraction of coefficient of variation of the genic variance among lines under strong stabilizing selection ($\Phi \to \infty$) as a function of $\Phi = N_e \varepsilon^2 V_E/[8(w^2+\sigma^2)]$. The curves are for a range of values of β, the shape parameter of the gamma distribution. All curves converge with increasing Φ as the CV becomes independent of the shape of the distribution but the shape parameter has increasing influence as $\Phi \to 0$.

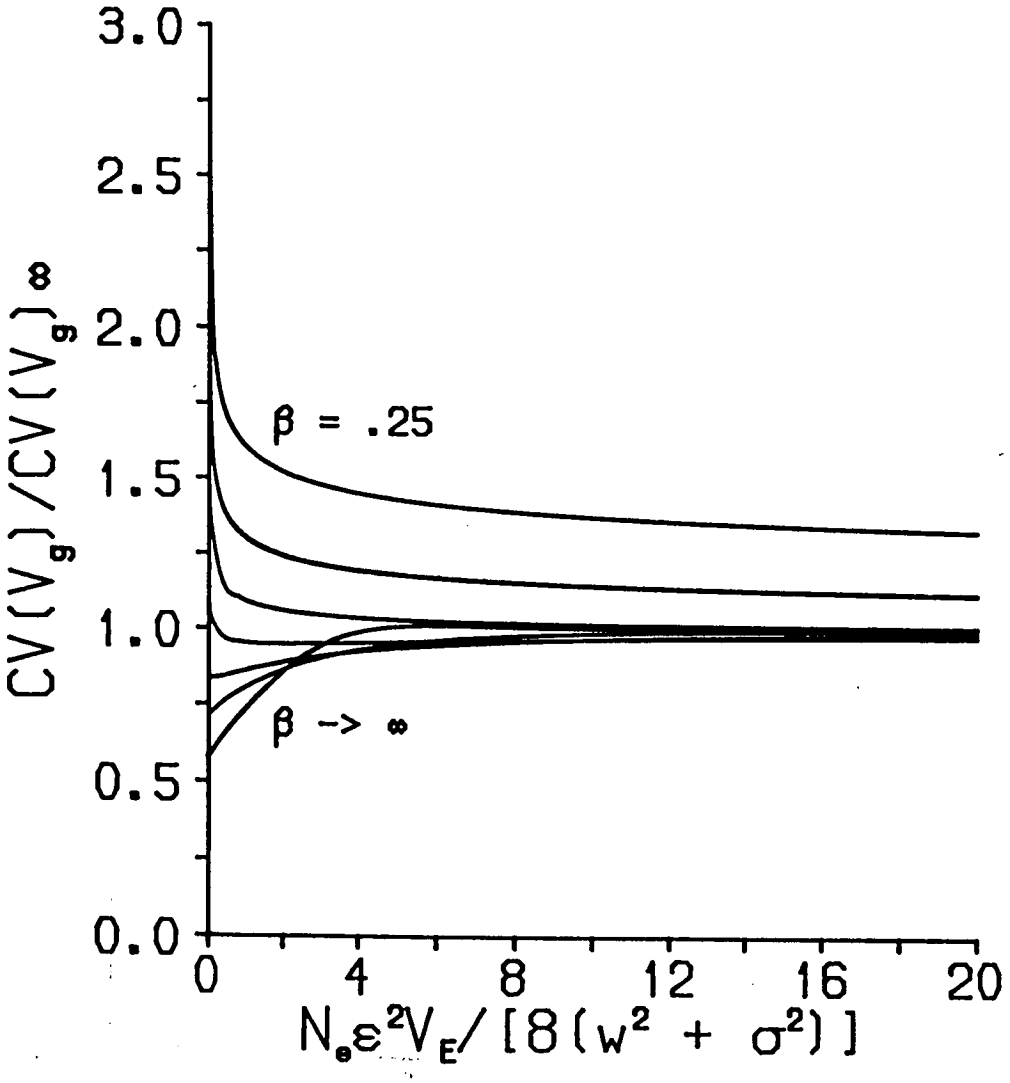


Fig. 1. The coefficient of variation of the genic variance among independent lines at equilibrium as a proportion of that predicted for very strong stabilizing selection, namely $[1/(4N_e\lambda)]^{1/2}$, as a function of $N_e E(s^*) = \Phi = N_e\varepsilon^2 V_E/[8(w^2 + \sigma^2)]$. The curves were generated by numerical integration of (5.5) and (5.7) using gamma distributions of mutant effects. The shape parameter ranges from $\beta \to \infty$ to $\beta = 1/4$ with intermediate values of $\beta = 8, 4, 2, 1, 1/2$.

5.3.2. Disequilibrium

The above analysis applies to the genic variance and its variance among independent lines. Such quantities cannot easily be measured. The additive variance and its variance among lines, which can be estimated, is influenced by departures from Hardy-Weinberg and linkage equilibrium at different loci affecting the trait. Bulmer (1976) and Avery and Hill (1977) showed that variation in disequilibrium can be an important contributor to the variation in the additive variation among lines. Here, Monte Carlo simulation is used to examine previous results on a neutral model and examine the effects of selection.

Neutrality. Using results obtained by Avery and Hill (1979), Lynch and Hill (1986) derived an expression for the coefficient of variation of the additive variance among independent lines at an equilibrium between drift and mutation in the absence of linkage,

$$CV(V_A) = [(E(a^4)/E^2(a^2))/(12N_e\lambda) + 2/(3N_e)]^{1/2}. \qquad (5.12)$$

(note, the additional n^{-1} term given by Lynch and Hill (1986) is inappropriate). The first term is the variance of the genic variance (cf. (5.9)) and the second is the variance of disequilibrium (note, although there is no selection and no net disequilibrium, the disequilibrium in each line varies stochastically about zero). The variance of the additive variance is therefore

$$V(V_A) = N_e\lambda E(a^4)/12 + 8N_e V_M^2/3. \qquad (5.13)$$

Bulmer (1980, Ch. 12) points out that there is an ambiguity in the interpretation of (5.12) and (5.13) because the variance of the additive variance depends on the number of individuals measured to estimate the additive variance within each line, and this may be different from the effective population size, N_e. As a starting point for resolving this difficulty, let N_e be the effective number of parents in each line (as before) and N' be the number of individuals per line used to estimate the additive variance. Assuming a normal distribution of observations, the variance in the estimate of the variance among lines due to

sampling is approximately $2V_A^2/(N' - 1) \simeq 8N_e^2 V_M^2/N'$. This additional source of variation affects $V(V_A)$ so (5.13) can be rewritten taking this source into account,

$$V(V_A) = N_e \lambda E(a^4)/12 + (2N_e V_M)^2[2/(3N_e) + 2/N'] . \tag{5.14}$$

This reduces to (5.13) as N' becomes large in which case the additional variance from estimating V_A becomes small. The results of Monte Carlo simulation of a neutral model with many alleles segregating are compared to evaluation of (5.14) in Fig. 2. Lines of various sizes were allowed to reach steady state and the additive variance and its variance among lines were computed using various numbers of progeny. The agreement between the models is very close.

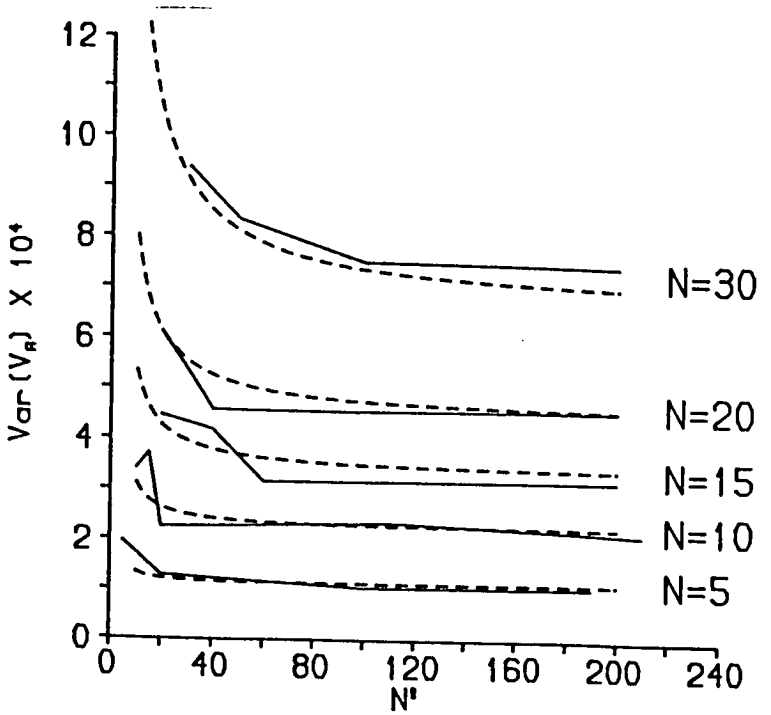


Fig. 2. The variance of the additive variance among independent lines for various parental population sizes (N) with variation in N', the number of individuals used to estimate V_A. The curves were generated by (a) Monte Carlo simulation (——) in the absence of selection and mutation parameters such that many alleles segregate ($\lambda = 0.2$, $\epsilon = 0.1$ sampled from a reflected gamma distribution with $\beta = 1/2$);
(b) equation (5.14) (----) with $E(a^4) = (\beta+2)(\beta+3)\epsilon^4 V_E^2/[\beta(\beta+1)]$ and $\beta = 1/2$.

There would be an additional source of variation in estimating $V(V_A)$ caused by error in estimating V_A within lines. The estimation might be done by, for example, offspring-parent regression or correlation of sibs.

For example, if V_A were estimated within lines from the covariance of half-sibs, it can be shown (cf. Robertson, 1959) that the variance of the estimate of V_A is approximately $2V_A^2[1 + 4/(\nu h^2)]^2/N'$, where ν is the number of progeny per half-sib family (assumed constant), h^2 is the heritability and N' in this case means the number of sires. Taking this additional source of variation into account, $V(V_A)$ (where the estimation of V_A is done by half-sib covariance) becomes,

$$V(V_A) = N_e \lambda E(a^4)/12 + (2N_e V_M)^2[2/(3N_e) + (1 + 8/(\nu h^2))^2/N'] \quad (5.15)$$

For small numbers of progeny and sires and traits of low heritability at equilibrium, the last term in (5.15) can dominate.

Stabilizing selection. No formulae are available to predict the variance among sublines of the disequilibrium component at steady state with stabilizing selection, although formulae for the expected disequilibrium with selection and an 'infinitesimal model' have been derived in Ch. 3. However, with free recombination, disequilibrium in a population has a very short 'memory' with on average half the previous disequilibrium lost due to recombination each generation. It is likely therefore that the additional term in (5.13) for the neutral case is a good predictor of the variation in disequilibrium for the case of stabilizing selection. Simulation runs (Table 5.2) show good agreement with eqn. (5.7) for $V(V_g)$ and (5.13) for $V(V_A)$.

Table 5.2

Comparison of predictions of variances of genic and additive variances among lines from Monte Carlo simulation and theory.

N_1	N_e	Theory $V(V_g)\times10^4$	Simulation $V(V_g)\pm1$s.e.$\times10^4$	Theory $V(V_A)\times10^4$	Simulation $V(V_A)\pm1$s.e.$\times10^4$
5	4.1	0.75	0.78±0.04	0.86	0.86±0.02
10	8.5	1.45	1.35±0.06	1.66	1.53±0.03
20	17.0	2.58	2.81±0.13	2.97	3.03±0.06
30	26.0	3.54	3.62±0.20	4.09	4.07±0.07

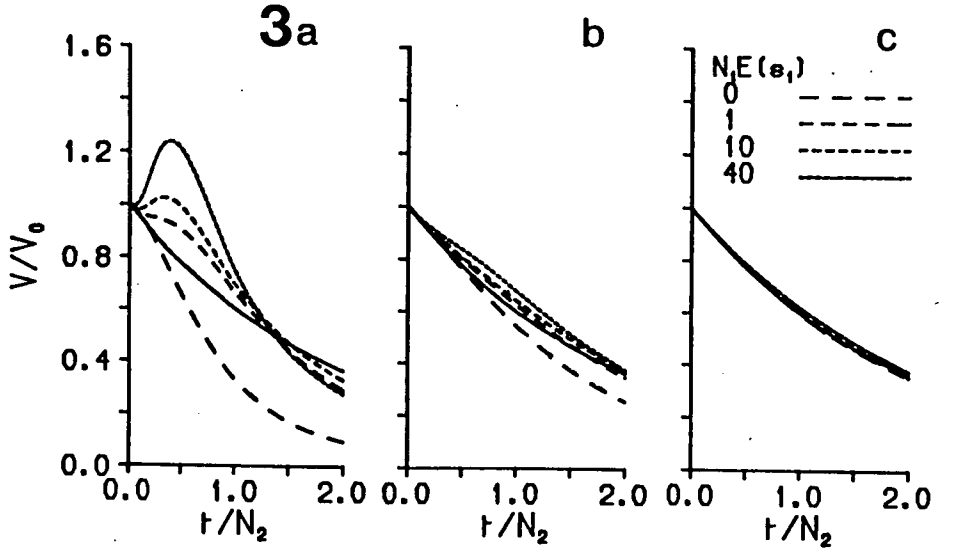
The mutation parameters of the Monte Carlo simulation were the same as Table 5.1 ($\lambda = 0.2$, $\varepsilon = 0.1$, and a gamma distribution of mutant effects with shape parameter $\beta = 1/2$ reflected about zero). The character was under stabilizing selection with $w^2 + \sigma^2 = 2$. The effective population size was measured in the simulation by following the fates of independent neutral alleles. The effective population size is less than the actual population size because of selection. The value of N_e computed by the computer program was used to compute the theoretical values in the Table. The theoretical value of $V(V_g)$ is from (5.7). The theoretical value of $V(V_A)$ is $V(V_g) + 2E(V_g)^2/(3N_e)$, i.e. includes the disequilibrium component from the neutral model (cf. (5.13)) and $E(V_g)$ was computed from (5.5).

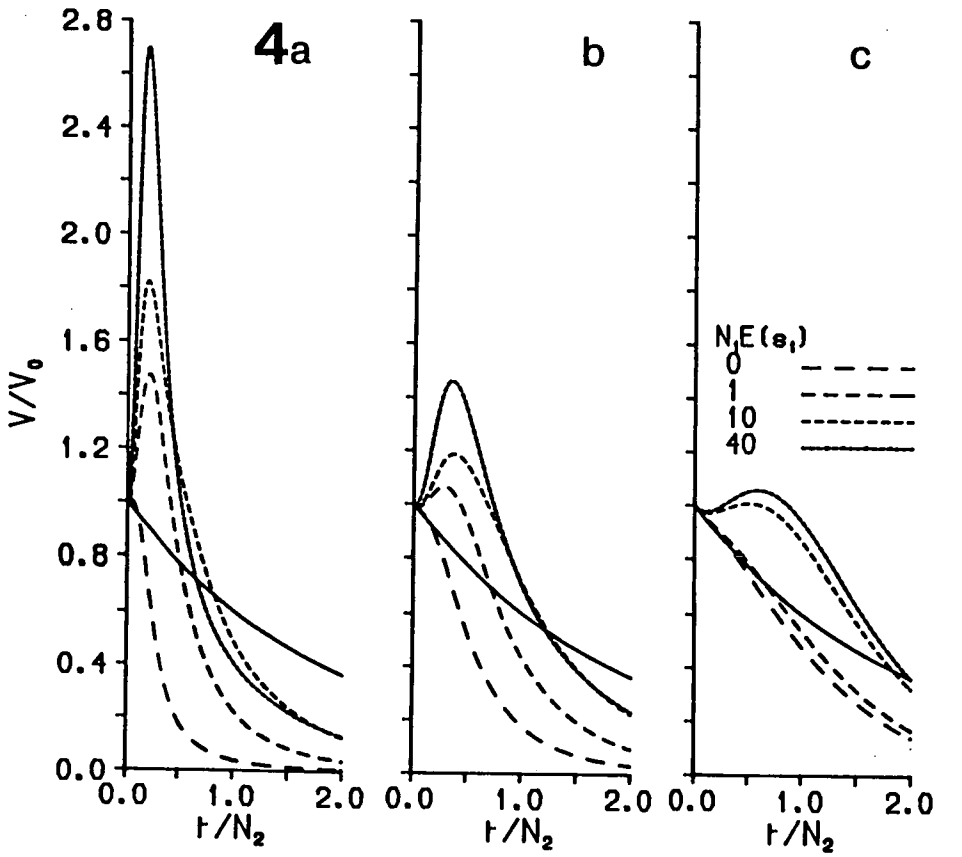
5.3.3. Effect of change of selection mode on variance

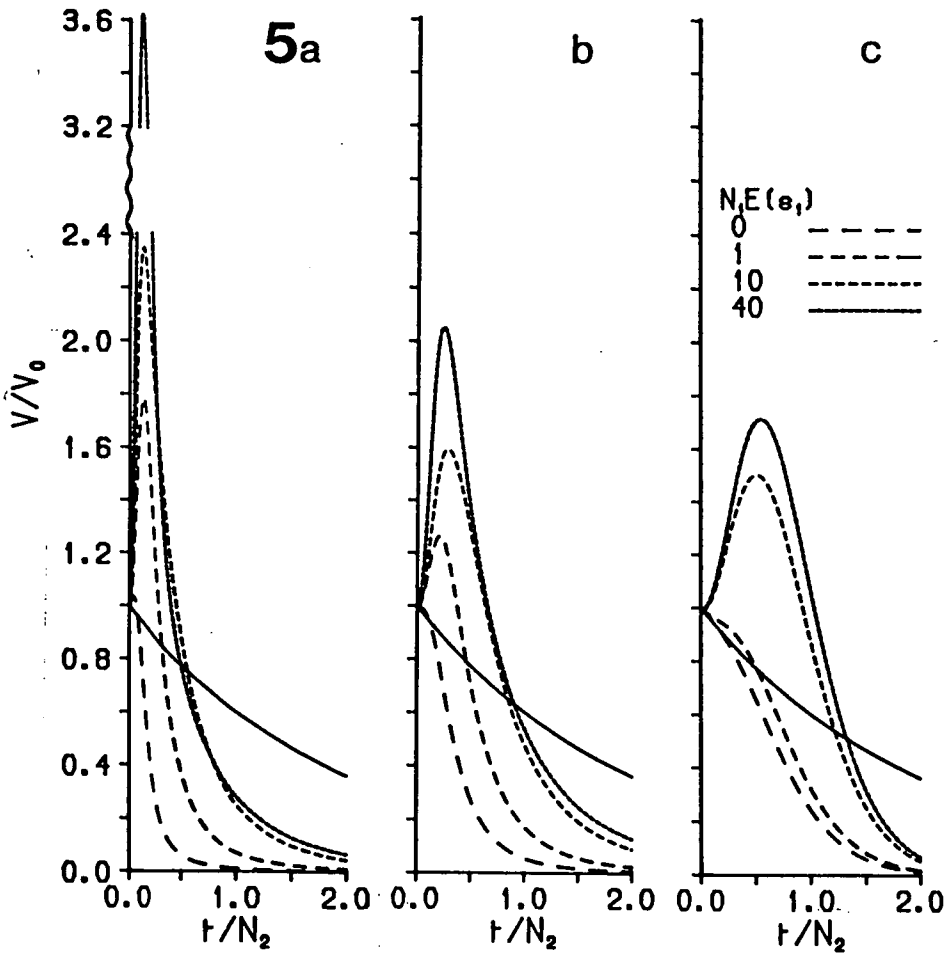
The results of the previous section show that the tendency for stabilizing selection to generate an extremely U-shaped distribution of allele frequencies influences the variance of the genetic variance between sublines. This effect also influences the pattern of response and change of variance of a character under stabilizing selection subsequently subjected to directional selection.

For a gamma distribution of mutant effects, the effect of selection on the genetic variance of the character is a function of three parameters: (1) $N_1E(s_1^*)$, the expected selective value in the stabilizing selection phase; (2) $N_2E(s_2)$, the expected selective value in the directional selection phase; (3) β, the shape

parameter of the gamma distribution. Figs. 3-5 show expected genic variances (ignoring disequilibrium) in the generations after a change in mode of selection for a range of $N_1 E(s_1^*)$ and $N_2 E(s_2)$.







Figs. 3–5. The expected genic variance as a proportion of that at $t = 0$ plotted against t/N_2 for four strengths of stabilizing selection, $N_1 E(s_1^*)$. Fig. 3. Strength of directional selection, $N_2 E(s_2) = 1$. Fig. 4. $N_2 E(s_2) = 5$. Fig. 5. $N_2 E(s_2) = 10$. The solid line on each Fig. is the expected variance for any value of $N_1 E(s_1^*)$ for no directional selection ($N_2 E(s_2) = 0$). Curves for three reflected gamma distributions of mutant effects are shown: (a) $\beta = 1/4$; (b) $\beta = 1$; (c) $\beta \to \infty$.

Mutants occurred only in the stabilizing selection phase and effects were sampled from reflected gamma distributions. Curves for three different values of the shape parameter, β, are shown: (1) β = 1/4, a highly leptokurtic distribution (see Ch. 2) (2) β = 1, an exponential distribution; (3) β → ∞, all mutant effects have equal absolute values. In all cases, the probability of a mutant of positive or negative effect was assumed to be the same. An implicit assumption of the analysis is that there is no stabilizing selection operating in the directional selection phase or, equivalently, that directional selection is strong relative to stabilizing selection. Figs. 3–5 show a wide range of values of the parameters. The selective values in the stabilizing selection phase range from $N_1 E(s_1^*) = 0$ (neutrality) to $N_1 E(s_1^*) = 40$. The latter case would pertain, for example, if $[E(a^2)]^{1/2} = 0.1$, $w^2 + \sigma^2 = 20$ and $N_1 = 6.4 \times 10^5$. The range of selective values during the directional selection phase is from neutrality ($N_2 E(s_2) = 0$), to $N_2 E(s_2) = 10$ (e.g. $E(|a|)/\sigma = 0.1$, $i = 1$, and $N_2 = 100$).

With values of $N_2 E(s_2)$ at the high end of the range shown, the pattern of change of variance departs substantially from that observed with neutrality. In some cases, there is a marked rise in variance followed by a rapid fall. The rise in variance occurs during the fixation of beneficial alleles segregating initially at low frequency. Such a pattern is therefore observed when all the following conditions pertain: (i) strong directional selection, so such alleles have a high probability of fixation; (ii) strong stabilizing selection because the probability distribution of allele frequencies becomes increasingly U-shaped with increasingly strong stabilizing selection, so the expected initial frequency of beneficial mutants of large effect is low; (iii) leptokurtic distribution of mutant effects (e.g. β = 1/4) because as the mutational distribution becomes more leptokurtic, a higher proportion of the mutational variance is contributed by mutations of large effect.

The pattern of rapid rise followed by rapid fall in variance depends on the presence of beneficial mutants in the directional selection phase, i.e. mutants of positive effect. In Figs. 3–5 there were equal probabilities of mutants of positive and negative effect, but the pattern of a rapid rise followed by a rapid fall in variance becomes more extreme with a higher proportion of mutants of positive effect (results not shown). In other cases, the variance falls off more quickly than for neutral genes. This occurs with weak stabilizing selection, but strong directional selection, in which case alleles initially segregating at intermediate frequencies become fixed at a high rate and the genetic variance

falls rapidly.

The curves show the expected genic variance and ignore the consequences of disequilibrium. Selection generates a negative disequilibrium component of variance which increases with increasingly tight linkage (see Bulmer, 1980 Ch. 9; Ch. 3). In such circumstances, the additive variance is less than the genic variance and the pattern of increase in additive variance would be less extreme than shown.

5.3.4. Variation in response.

Using similar methods to the above, Hill and Rasbash (1986) analysed the variation in response to directional selection. Higher variation in response was noted with increasingly leptokurtic distributions of effects of segregating alleles and with increasingly U-shaped probability distributions of allele frequency. The variation in the genic variance or response can be easily computed with a transition matrix. Using Monte Carlo simulation, however, it is possible to generate replicates of responses for different $N_1 E(s_1^*)$ and $N_2 E(s_2)$ parameter combinations and the general pattern of the response is perhaps easier to visualize (and compare to the results of experiments).

Table 5.3 shows cumulative responses and CV's of cumulative responses to generations 10 and 20 among 10 independent replicates sampled from independent populations initially at a mutation-stabilizing selection balance with a range of strengths of stabilizing selection and sizes of gene effects (examples of responses are plotted in Figs. 6 and 7).

Table 5.3

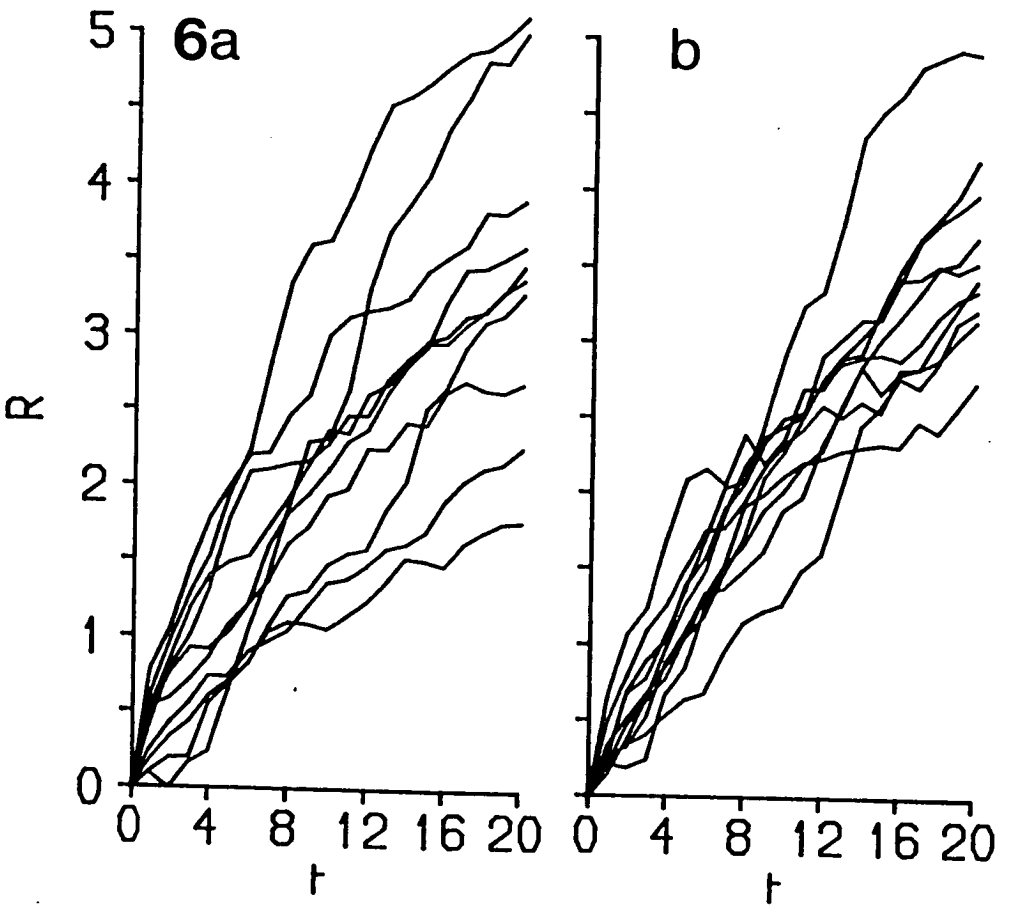
The mean, coefficient of variation (CV), and range of cumulative response to selection to generations t = 10 and t = 20 from 10 independent populations initially at a mutation–stabilizing selection–drift equilibrium.

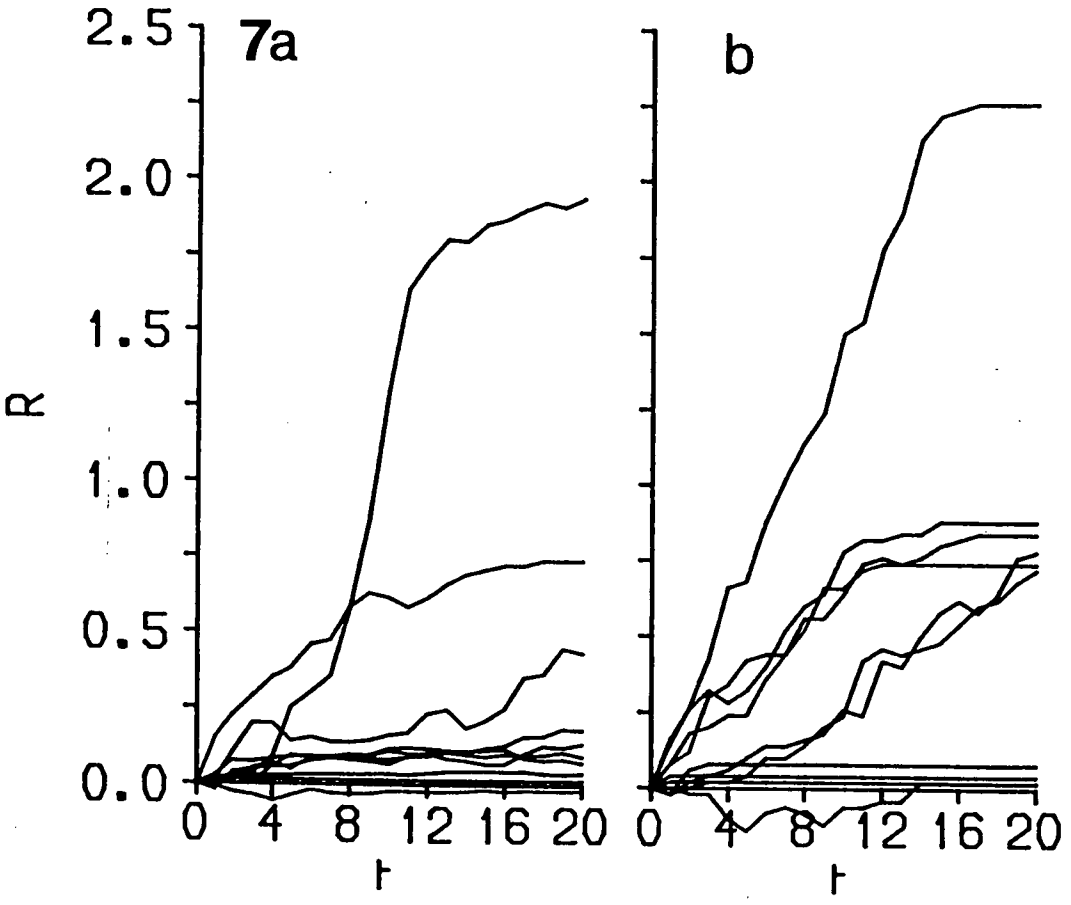
$N_1 E(s_1^*)$	ε	β	t	Mean	CV	Range
0	0.2	∞	10	2.21	0.19	1.27 - 2.90
			20	3.62	0.17	2.74 - 4.91
0	0.2	½	10	2.15	0.36	1.06 - 3.63
			20	3.35	0.34	1.77 - 5.12
0	0.8	∞	10	1.69	0.44	0.52 - 2.82
			20	2.13	0.41	0.88 - 3.34
0	0.8	½	10	1.78	0.51	0.39 - 3.38
			20	1.49	0.66	0.39 - 3.45
½	0.2	∞	10	1.36	0.32	0.64 - 2.29
			20	2.46	0.22	1.85 - 3.82
½	0.2	½	10	0.66	0.36	0.42 - 1.18
			20	1.16	0.41	0.58 - 1.95
8	0.8	∞	10	0.42	1.14	0.06 - 1.50
			20	0.64	1.06	0.02 - 2.26
8	0.8	½	10	0.25	1.64	0.02 - 1.31
			20	0.36	1.67	0.02 - 1.94

Mutations occurred in the stabilizing selection phase only and were sampled from a reflected gamma distribution with shape parameter β and scale parameter ε given in the table. The mutation rate λ was such that $V_M / V_E = 10^{-3}$. The population size in the stabilizing selection phase was $N_1 = 160$ and in the directional selection phase was $N_2 = 20$.

The Table compares results from a reflected gamma distribution of mutant effects with shape parameter $\beta = 1/2$ and equal probabilities of positive and negative effects ($\beta \to \infty$). The population size in the stabilizing selection phase was $N_1 = 160$ and in the directional selection phase was $N_2 = 20$, and V_M was 10^{-3}. The main points to note from the table are: (i) In theory, the average initial response rate is equal to the standing additive variance in the stabilizing selection phase. With no stabilizing selection ($N_1 E(s_1^*) = 0$), the theoretical initial rate is therefore $2 N_1 V_M = 0.32 V_E$, but the cumulative response

to generation 10 was less than ten times this because of the presence of disequilibrium generated by directional selection and a loss of genetic variance due to changes in gene frequency; the average response with stabilizing selection is lower than the average response from initially unselected populations because the expected steady state variance is lowered; (ii) the more leptokurtic distribution of mutant effects ($\beta = 1/2$) gives a higher coefficient of variation of response than that for equal absolute values of mutant effects; (iii) with stronger stabilizing selection ($N_1 E(s_1^*) = 8$), because there is little standing variance, average response $\bar{z}$ generally small, but occasionally an allele of large effect segregating at low frequency gives a rapid early response so the range of response is large relative to the mean; (iv) many response patterns give similar means, variances and ranges of response and in practice would not be distinguishable from one another. Some examples of responses, the results of which are summarized in Table 5.3, are shown in Figs. 6 and 7.





Figs. 6-7. Examples of selection responses generated by Monte Carlo simulation from directional selection of samples of N_2 = 20 individuals from independent populations of size N_1 = 160 individuals at mutation-stabilizing selection-drift equilibrium. Fig. 6. There was no selection in the stabilizing phase. The sizes of gene effects were given by ε = 0.2 and V_M was 10^{-3}. Fig. 7. The strength of stabilizing selection was such that $N_2 E(s_2)$ = 8, and the value of ε was 0.8 with $V_M = 10^{-3}$. (a) Reflected gamma distribution with β = 1/2. (b) Equal probabilities of positive and negative effects.

Figs. 6a and 6b show results for a reflected gamma distribution with shape parameters $\beta \to \infty$ and $\beta = 1/2$ respectively for the case of no stabilizing selection and small gene effects ($\varepsilon = 0.2$). These response patterns are similar to those commonly observed in selection experiments, although it should be emphasized that weak stabilizing selection would not lead to much change in pattern. In contrast, Figs. 7a and 7b which again are for cases of reflected gamma distributions with shape parameter $\beta = \infty$ and $\beta = 1/2$ respectively show much higher variation in response. In these cases, stabilizing selection was relatively strong ($N_1 E(s_1^*) = 8$) and gene effects were large ($\varepsilon = 0.8$). Such responses are not typical of selection experiments.

The responses in Figs. 6 and 7 show the change in mean genotype, so there would be more variation in mean phenotype than shown because of the presence of an environmental component of variation. However, this would contribute little to the variation of response of response unless N_2 is very small or the heritability of the character is low. The simulated selection responses were generated from independent populations. In practice, a caged population is often used to initiate independent lines, so variation in genes segregating initially and hence variation in response would be due to sampling from the base population rather than to different genes segregating in independent populations. Simulation showed that responses generated from sub-populations of single rather than a set of independent populations tend to vary less than the responses discussed above because of reduced variance between lines in the initial genetic variance and alleles of large effect which are not rare are likely to be fixed in all replicate lines. If, however, the mutational variance is generated by few mutants of large effects (e.g. $\varepsilon = 0.8$, cf. Fig. 7), under stabilizing selection, because genes segregate at very low frequency at equilibrium, the variation in response is very similar to that obtained by sampling from independent populations.

5.4. Discussion

5.4.1. Mutation-Stabilizing Selection Balance

The two allele models of Latter (1960) and Bulmer (1972) and Turelli's (1984) "House of Cards" model showed that the expected genic variance in an infinite population is independent of the effects of alleles at the loci controlling the trait. Similarly, the results show that with strong stabilizing selection the

variance of the genic variance among independent lines is also independent of the size of gene effects, and is only a function of the effective population size, strength of stabilizing selection, and the genomic mutation rate, λ. As drift becomes more important relative to selection, the coefficient of variation of the genic variance becomes increasingly dependent on the shape of the distribution of effects of new mutants, but we have little if any information about this parameter and can only conjecture that distributions of mutant effects are very leptokurtic (Robertson, 1967; Shrimpton and Robertson, 1988; Ch. 4). However, for the reflected gamma distribution with shape parameter $\beta = 1/4$, from which some mutants are of very large effect indeed, the coefficient of variation of V_g is relatively independent of the strength of stabilizing selection and over a wide range of values of the relevant parameters. Variation in estimates of genetic variance from different populations therefore does not necessarily tell us much about the selective forces operating in the population.

Turelli's (1984) "House of Cards" analysis of mutation-stabilizing selection balance is multi-allele, but the formula for the expected genic variance at equilibrium is the same as obtained from the two allele analyses of Latter (1960) and Bulmer (1972). Why is this so? In these models, the population size is assumed to be infinite, or equivalently, selection is assumed to be very strong. The probability distribution of allele frequencies is therefore very U-shaped with alleles at intermediate frequencies absent. Each new mutant allele almost always occurs at a locus previously carrying the 'wild type' allele. The fates of new mutant alleles are therefore almost independent of any other mutant alleles segregating at the same locus in the population, and a two allele analysis with the parameter $n\mu$ replaced with genomic mutation rate, λ, is sufficient. Similarly, with weak selection the fates of different alleles at the same locus are essentially independent of one another, they can be considered as occurring at separate loci, and the two allele treatment also applies. As shown in Ch. 4, the model of stabilizing selection is very similar to a model of unconditionally deleterious genes with selection coefficient s proportional to α^2 and independent of gene frequency, q.

The tendency for stabilizing selection to generate an extremely U-shaped probability distribution of allele frequencies has important consequences for subsequent changes of variance and hence responses with directional selection. Under certain circumstances, namely strong directional and strong

stabilizing selection, large increases in variance occur in early generations due to the fixation of alleles initially at low frequency; in other circumstances, namely strong directional selection and weak stabilizing selection, a rapid fall in variance (much faster than expected from drift alone) can occur due to the rapid fixation of genes initially at intermediate frequencies. Seldom, if ever, are such response patterns seen in selection experiments. For example, the replicated Drosophila abdominal bristle selection experiments of Clayton, Morris and Robertson (1957), Frankham, Jones and Barker (1968) and Yoo (1980), which were initiated from cage populations, showed little sign of early accelerated or rapidly falling responses. Similar regular patterns were observed in a Tribolium egg production selection experiment (i.e. a character closely related to fitness) (Ruano, Orozco and Lopez-Fanjul, 1975) initiated from a cage population, and in a selection experiment of cannon-bone length in Scottish blackface sheep (Atkins and Thompson, 1986). In the latter case, Atkins and Thompson showed that the response closely matched the predicted response of an 'infinitesimal model' which incorporated the effect of disequilibrium on the additive variance (Bulmer, 1980 Ch. 9). The results of the experiment of Frankham, Jones and Barker (1968) are of particular interest because selection on bristle score was performed using various population sizes and a range of selection intensities. The expected response can be estimated if the 'infinitesimal model' is assumed and the initial genetic variance is obtained from: (i) the heritability estimated from the base population; or (ii) the realised heritability estimated from the selection response in the first one or two generations which in theory is almost independent of the magnitude of gene effects. Such analysis shows substantially lower predicted responses using the base population genetic variance estimate than observed in the experiment. Using the genetic variance obtained from the realised heritability in the first two generations, the agreement between the experimental results and the infinitesimal model is closer. With the strongest selection strength (10%) and the biggest population (N = 80), some hint of an accelerating response was observed, but unfortunately this line was not replicated. The responses from Yoo's (1980) long-term experiment fit closely an infinitesimal model if parameters derived from the initial generations are assumed although the response continued longer than predicted by the infinitesimal model. The results of Falconer's (1973) replicated mouse body weight selection experiment also also give a reasonable fit to the infinitesimal model if the genetic variance derived from the response in the initial generations is assumed, but this realised heritability

is rather higher than Falconer's estimate of the heritability in the base population. In this case the lines were derived from crosses of inbred lines and presumably some alleles were initially at intermediate frequencies.

Surprisingly, the experimental selection response patterns do not tell us very much about the strength of stabilizing selection which might affect the characters in the base population because the expected variance each generation of directional selection becomes the same irrespective of the distribution of allele frequencies as $N_2 s_2 \to 0$. They indicate, however, that selective values of directional selection ($N_2 E(s_2)$) cannot in general be very high, say greater than one, because either a rapid rise or a rapid fall in variance would have been observed in some cases.

Many of the responses generated by Monte Carlo simulation, using a wide range of parameters, are very similar to selection responses obtained experimentally. Some types of response patterns in Table 5.3 and Figs. 6–7 are not, however, observed experimentally. For example, with $\varepsilon = 0.8$ and a gamma distribution with shape parameter $\beta = 1/2$, responses are very variable because there are few genes segregating. This pattern becomes more extreme with strong stabilizing selection ($N_1 E(s_1^*) = 8$, Fig. 7b) with some lines giving a rapid early burst of response. Bursts of response have been seen in selection lines, generally in long-term experiments (Thoday, Gibson and Spickett, 1964; Yoo, 1980), and are most likely the result of fixation of mutants appearing since the start of the experiment (Hill, 1982b). The alternative hypothesis of segregation of rare recessive alleles is also possible, though unlikely if the burst occurs late in the experiment as in the cases cited above (Robertson, 1978). Breakdown of linkage disequilibrium is also an unlikely explanation (Keightley and Hill, 1983).

5.4.2. Concluding Remarks

The validity of the stabilizing selection model of natural selection has been discussed extensively elsewhere (Robertson, 1973; Turelli, 1984, 1985; Ch. 4). The most important weakness is that the pure stabilizing selection model ignores selection which might be acting at the locus through pleiotropic effects on characters directly related to fitness. Other models where the mutant allele is at a selective disadvantage (e.g. Hill and Keightley, 1988) have similar qualitative effects on the probability distribution of allele frequencies. The

essential problem in explaining quantitative genetic variation is not whether mutation is an adequate force to explain observed variation, for in the absence of selection it is a more than adequate force. The problem is the mode of action of natural selection and the selective values of the genes affecting the character.

The analysis here is purely additive and ignoring dominance is a serious limitation. As shown by Kacser and Burns (1981), the larger the absolute effect of a mutant allele, the more likely it is to behave as nearly recessive. The consequences of this could be deduced with specific models of the relationship between mutant effect and dominance. It is likely that the tendency to give burst or rapid falls in response would be reduced, however, because alleles of large additive effect would contribute little to the variance in the stabilizing selection phase and would have little chance of fixation from directional selection.

Other models of the mutation process might also be considered. For example, Cockerham and Tachida's (1987) model differs from the present step-wise model because the effect of a new mutation replaces the current value at the locus, not as in this case adding to the value. This additional constraint does not affect the equilibrium behaviour with stabilizing selection as the model is formally the same as the "House of Cards". It can lead, however, to limits in the case of directional selection. The present results, therefore, would only be applicable in the short term.

Chapter 6
Models of Quantitative Variation of Flux in Metabolic Pathways

6.1. Introduction

The genetic variation of any character must be the result of genetic variation affecting basic biochemical processes. To what extent can advances in understanding biochemical systems improve our knowledge of the genetics of quantitative variation? Enzymology and studies of intermediary metabolism have provided insights into the actions of individual enzymes and the coupling of reactions in complex pathways represented graphically as the 'metabolic map'. A different approach, metabolic control theory (see review by Kacser and Porteous, 1987), examines the effects of changes in enzyme activity on metabolic fluxes and metabolite concentrations. An analogous, but more complex theoretical treatment of control of metabolic systems has been developed by Savageau and collaborators (Savageau, 1969). (See Savageau, Voit and Irvine (1987) for a discussion of the relationship between the approaches). Fluxes and metabolite concentrations are known as *variables* in the system while enzyme activities are under genetic control and are known as *parameters*. In metabolic control theory, the importance of an enzyme in controlling a flux is quantified by a 'control coefficient' (Kacser and Burns, 1973), the fractional change in flux resulting from a small fractional change in enzyme activity. The effects of discrete changes in enzyme activity are important, however, because mutants can cause large changes. The problem of the general recessivity of 'null' mutants at enzyme loci has been investigated by Kacser and Burns (1981) who offered a general explanation based on the properties of metabolic pathways.

The effects of enzyme activity variation can be viewed at several levels. At the first, there is experimental evidence that enzyme activity itself is polygenically determined (e.g. Paigen, 1979; Laurie-ahlberg *et al.*, 1980). At the next level are metabolic fluxes and metabolite pool concentrations. These are functions of perhaps all the enzymes in the 'metabolic map', but are presumably more closely related to those enzymes in the pathway which 'carries' the flux. At the third level are characters which are traditionally the subject of quantitative genetics, some of which have economic importance (e.g. growth rate, fatness), and likely to be influenced by many metabolic fluxes. Here, quantitative

variation in metabolic flux is modelled by assuming that the phenotype (flux) is affected by enzyme activity variation segregating in a population. The flux is assumed to be a function of enzyme activity according to a simple biochemical model. The variance in flux is partitioned into additive and non-additive components. This approach differs from classical models in quantitative genetics where the phenotypic variation usually depends on arbitrary assumptions of the effects of alleles. Here, the variation in the phenotype resulting from genetic variation depends on biochemical assumptions. It is known that because enzymes share common substrates, allelic substitutions at one locus modify the allelic effect at another (e.g. Dykhuizen, Dean and Hartl, 1987), so epistatic interactions are automatically built into metabolic systems. The conditions necessary for substantial interaction variance in flux as a proportion of the total genetic variance are investigated (the terms interaction and epistasis are used synonymously here). An attempt is made to relate experimentally determined variations in enzyme activities to the predictions of the models. Most of the formulations and simulations are restricted to the case of haploid populations. The consequences of diploidy and a dominance component of variance are discussed, however, in relation to results of previous investigations (Kacser and Burns, 1981).

6.2. Model

6.2.1. Biochemical model

Consider the following monomolecular transformation within a pathway of substrate S_i to product S_j catalysed by enzyme E_j.

$$\ldots \rightleftharpoons S_i \overset{E_j}{\rightleftharpoons} S_j \rightleftharpoons \ldots$$

The rate, v_j of the reaction is given by

$$v_j = (V_j/M_i)(S_i - S_j/K_{i,j})/(1 + S_i/M_i + S_j/M_j) \qquad (6.1)$$

(Cleland, 1963), where V_j is the maximal velocity (V_{MAX}), M_i is the Michaelis

constant (K_m) of E_j with respect to S_i, similarly M_j is the Michaelis constant for E_j with respect to S_j, S_i and S_j are the concentrations of S_i and S_j respectively, and $K_{i,j}$ is the equilibrium constant for the step. Assume that the enzyme is acting in conditions of low saturation, that is, the Michaelis constants for the forward and backward reactions are substantially greater than S_i and S_j respectively. An approximate expression for the rate is therefore

$$v_j = (V_j/M_i)(S_i - S_j/K_{i,j}). \tag{6.2}$$

The quantitative variation in the character metabolic flux is investigated in two types of pathways containing a number of such enzymes:

Linear Pathway. Fig. 6.1. The external substances X_A (input) and X_B (output) are assumed to be at constant concentrations determined by the external environment. This would be the case, for example, if X_A were a nutrient supplied to the organism at a constant rate, similarly if X_B were a waste or end product, changes in whose concentration due to the activity of the pathway can be ignored.

$$X_A \underset{E_1}{\rightleftharpoons} S_1 \underset{E_2}{\rightleftharpoons} S_2 \underset{E_3}{\rightleftharpoons} \ldots S_{n-1} \underset{E_n}{\rightleftharpoons} X_B$$

Fig. 6.1. The linear pathway. A chain of n monomolecular unsaturated enzymes catalyses the conversion of externally controlled substrate X_A to product X_B via the intermediates S_1 to S_{n-1}.

The system will reach a steady state at which all substrate concentrations are constant (the steady state concentrations) and the rates of all the steps in the pathway are equal to one another and hence to the pathway flux, J, the rate of consumption of X_A and production of X_B. Expressions of type (6.2) may be written down for each step in the pathway which are a consistent set of simultaneous equations. Elimination of the variable S terms gives an expression for the flux in terms of parameters only,

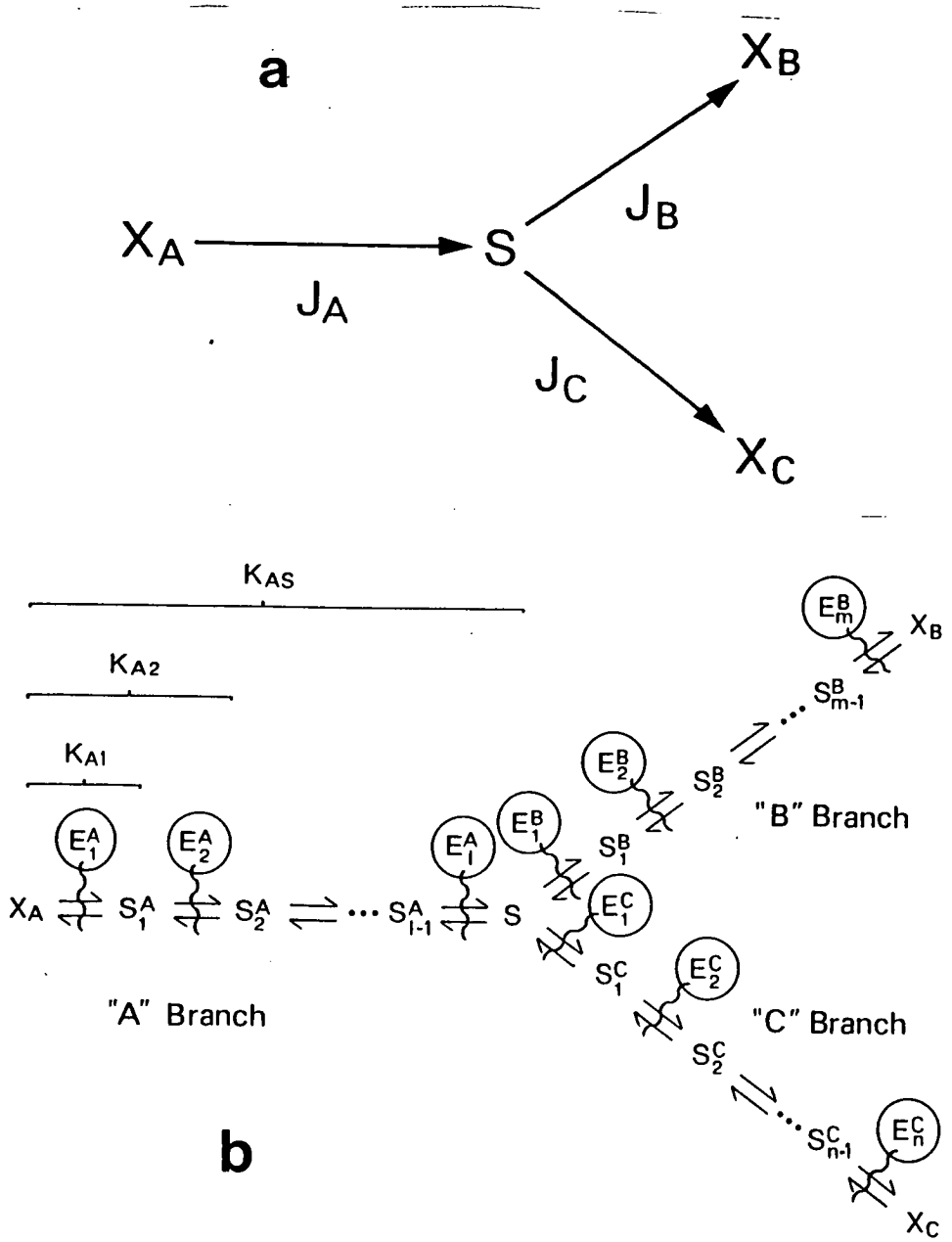
$$J = (X_A - X_B/K_{A,B}) \; / \; (\Sigma_{i=1}^{n} M_{i-1}/V_i K_{A,i-1}) \qquad (6.3)$$

(Kacser and Burns, 1973), where X_A and X_B are the concentrations of X_A and X_B respectively, $K_{A,B} = K_{A,1} \cdot K_{1,2} \cdots K_{n-1,B}$ is the equilibrium constant between X_A and X_B, similarly $K_{A,i-1}$ is the equilibrium constant between X_A and S_{i-1}, the first term in the summation containing $K_{A,0} = 1$, M_0 is the Michaelis constant of E_1 with respect to X_A, and n is the number of steps in the pathway. Equation (6.3) may be rewritten as

$$J = (X_A - X_B/K_{A,B}) \; / \; (\Sigma_{i=1}^{n} 1/E_i). \qquad (6.4)$$

The E terms are proportional to the enzyme activities which are subject to genetic variation, and are also functions of equilibrium constants.

Branched Pathway. Fig. 6.2. As in the linear pathway, the concentrations of the external substances, in this case X_A, X_B and X_C, are assumed to be constant and determined by the environment.



Figs. 6.2a and 6.2b. The branched system. The system is bounded by the external substances X_A, X_B and X_C. The three symmetrical branches are labelled "A", "B" and "C". The point of divergence in the system is at substrate S which is catalysed by three enzymes E^A_1, E^B_1 and E^C_1. The superscript on the enzyme and substrate symbols refer to the branch, and the subscript refers to the position in the branch. Equilibrium constants are similarly named. For example, $K_{A,1}$ refers to the equilibrium constant of the reaction $X_A \rightarrow S^A_1$ etc. For the unsaturated case, all the enzymes in each branch can be condensed to 'group enzyme activities', *a*, *b* and *c*.

There are three branches, and the following terminology is used to describe them: (i) The *common* branch is "A"; (ii) The *reference* branch is "B"; (iii) The *competing* branch is "C". At steady state, the flux in the common branch is the sum of the fluxes in the reference and competing branches, i.e.

$$J_A = J_B + J_C. \tag{6.5}$$

The substrate S is catalysed by the first enzyme in both the reference and competing branches which therefore compete for S. At steady state the rates of each step in the system can be described by equations of type (6.4) which equate to a branch flux. Together with relation (6.5), elimination of the variable S from these equations yields an expression for the flux in the reference branch:

$$J_B = [c(X_A - X_B/K_{A,B}) + a(X_C/K_{S,C} - X_B/K_{S,B})] \,/\, (ab + ac + bc/K_{A,S}) \tag{6.6}$$

where the terms a, b, and c are 'group enzyme activities', and $K_{A,B}$, $K_{S,B}$, $K_{S,C}$ and $K_{A,S}$ are equilibrium constants between X_A and X_B, S and X_B, S and X_C, and X_A and S respectively. For example, the a group enzyme activity is given by

$$a = \Sigma_{i=1}^{l} 1/E_i^A,$$

where l is the number of enzymes in the "A" (common) branch, and the E terms are proportional to enzyme activity and have similar meanings as in (6.4). The b and c group activities are given by $b = \Sigma 1/E_i^B$ and $c = \Sigma 1/E_i^C$.

6.2.2. Genetic Model

Basic Assumptions. It is assumed that genetic variation affects enzyme activities in a population of haploid organisms. Each activity is assumed to be affected by variation at one locus only and there are up to two alleles at each locus. The allele of higher activity at locus j is termed the *wild type* , E_{1j}, and the lower activity allele is termed the *mutant*, E_{2j}. The frequency of the wild type allele at the locus is q_j. Linkage disequilibrium is assumed to be absent.

Variance Components of Flux. Let n be the number of loci affecting the flux and g be the number of different genotypes, which in a haploid population is 2^n. Let J_i be the value (i.e. flux) of genotype i and f_i be its frequency in the population. The total genetic variance is

$$V_T = \Sigma_{i=1}^{g}(J_i - m)^2 f_i, \tag{6.7}$$

where $m = \Sigma J_i f_i$ is the population mean. The additive variance in flux is

$$V_A = \Sigma_{j=1}^{n}\alpha_j^2 q_j(1 - q_j), \tag{6.8}$$

where α_j is the average effect of a gene substitution (Falconer, 1981 Ch. 7), which is the mean change in value resulting from the change of a mutant allele to a wild type allele, and can be obtained by differentiating the mean flux with respect to gene frequency, q_j (see Kojima, 1959). The resulting expression for α_j is,

$$\alpha_j = \partial m/\partial q_j = \Sigma_{i=1}^{g} f_i J_i(\delta_{ij}/q_j - (1 - \delta_{ij})/(1 - q_j)), \tag{6.9}$$

where δ_{ij} is one if the genotype is wild type with respect to locus j, or zero if mutant. Substitution of an expression for flux (e.g. (6.4) or (6.6)) gives an equation for the variance in terms of biochemical parameters (e.g. enzyme activities). The interaction component is simply

$$V_I = V_T - V_A. \tag{6.10}$$

6.3. Results

6.3.1. Interaction in the Linear Chain

Two Locus Interactions. Assume that allelic variation in enzyme activity occurs at loci j and k. With two variable loci there are therefore four haploid genotypes. Let the value of these be J_{11}, J_{12}, J_{21}, and J_{22} where, for example,

J_{12} is the value of the flux for the wild type at locus j and the mutant at locus k. This is

$$J_{12} = (X_A - X_B/K_{A,B})/(1/E_{1j} + 1/E_{2k} + K), \qquad (6.11)$$

where K is the sum of reciprocals of enzyme activities other than j and k as in (6.4). Similar expressions can, of course, be written down for the other three genotypes. An explicit expression for the additive-by-additive interaction variance, V_I can be obtained using Kojima's (1959) method of differentiating the mean genotypic value with respect to the gene frequencies, q_j and q_k. This is

$$\begin{aligned} V_I &= q_j(1 - q_j)q_k(1 - q_k)[\partial^2 m/(\partial q_j \partial q_k)]^2 \qquad (6.12) \\ &= q_j(1 - q_j)q_k(1 - q_k)\varepsilon^2, \end{aligned}$$

where $\varepsilon = J_{11} - J_{12} - J_{21} + J_{22}$, the interaction term for the two locus case. Consider the ratio of the additive to interaction variance, V_A/V_I. From (6.8) and (6.12) and after rearrangement, this is

$$V_A/V_I = (a_1/\varepsilon + q_k - 1)^2/[q_k(1 - q_k)] + (a_2/\varepsilon + q_j - 1)^2/[q_j(1 - q_j)], \qquad (6.13)$$

where $a_1 = J_{11} - J_{21}$ and similarly $a_2 = J_{11} - J_{12}$. The ratio of variances as a function of gene frequencies and enzyme activities can be obtained by substituting expressions of type (6.11) for the flux into (6.13). This has the following properties.

(i) V_A/V_I is independent of the environmentally controlled parameter $X_A - X_B/K_{A,B}$.

(ii) The values of the terms a_1/ε and a_2/ε are in the range one to infinity because both have the form $x/(x - y)$, $0 < y < x$. A very large value of either a_1/ε or a_2/ε implies very little interaction variance relative to additive variance for all q_j and q_k. It can easily be shown that this occurs as $E_{2j}/E_{1j} \to 1$ and/or

$E_{2k}/E_{1k} \rightarrow 1$. The interaction variance is highest relative to the additive variance as $a_1/\varepsilon \rightarrow 1$ and $a_2/\varepsilon \rightarrow 1$, i.e. when $J_{22} - J_{12}$ and $J_{22} - J_{21}$ become small. This occurs as the activities of the mutant alleles approach zero (null alleles).

(iii) With a small interaction term ($\varepsilon \rightarrow 0$), V_A/V_I is at a maximum as a function of gene frequency when q_j and q_k are 1/2.

(iv) With null alleles at both loci (which gives maximum epistasis, see (ii) above), the interaction variance is highest relative to the additive variance as $q_j \rightarrow 0$ and $q_k \rightarrow 0$ (i.e. the mutant alleles are common).

(v) The activities of the non-variable enzymes in the pathway are represented in the constant K. As $K \rightarrow \infty$ and if the mutant activities are not null, V_A/V_I becomes very large because $a_1/\varepsilon \rightarrow \infty$ and $a_2/\varepsilon \rightarrow \infty$. This implies that the interaction variance is small relative to the additive variance if most of the control of flux lies with non-variable loci (see Kacser and Burns, 1973, 1981).

Multiple Locus Interactions. Interactions between pairs of loci are of lowest order, but multiple locus interactions also contribute to the genetic variance. Using Kojima's (1959) method, it is possible to obtain expressions for the variance contributed by higher order interactions, but these are very complex and the general properties are better illustrated by evaluating particular parameter sets on a computer. Fig. 6.3 shows results for a linear chain of ten equivalent loci for a range of values of the ratio of mutant to wild type activity, E_2/E_1. Results for three gene frequencies of the wild type are shown (q = 0.1, mutant allele common; q = 0.5; q = 0.9, mutant allele rare), and are expressed as the proportion of interaction variance, V_I/V_T. Each step on the horizontal axis represents changing one non-variable locus in the chain to a variable locus of given gene frequency in the population.

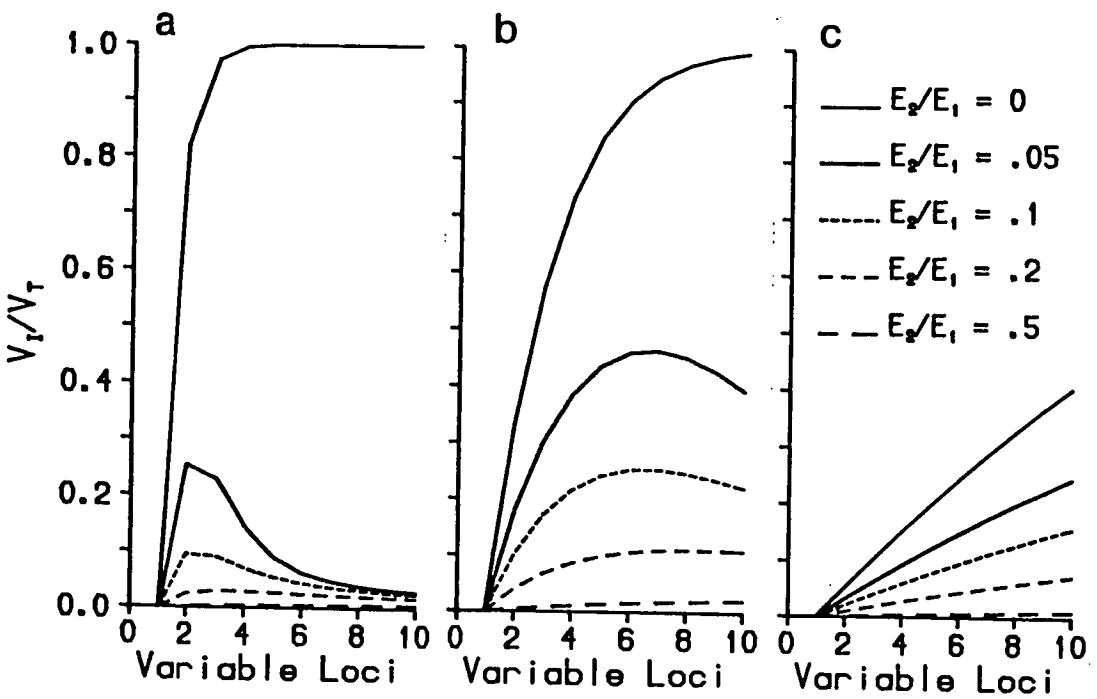


Fig. 6.3. The proportion of interaction variance of flux in a fixed length pathway of 10 equivalent loci. Each step on the horizontal axis corresponds to the introduction of one further variable locus into the population. The curves relate various values of ratio of mutant to wild type activity, E_2/E_1. (a) q = 0.1; (b) q = 0.5; (c) q = 0.9.

As indicated by (6.13), with two variable loci only, the proportion of interaction variance is highest for null mutant alleles when the mutant allele is common (q = 0.1 in Fig. 6.3). The presence of further variable loci does not necessarily lead to an increase in the proportion of interaction variance. Why is this so? Although an increase in the number of variable loci leads to an increase in the number of interaction terms, increasing the number of mutant alleles segregating lowers the average activities of the enzymes in the chain. This has the consequence that the average change in flux between mutant and wild type becomes more linear, hence the presence of maxima for all cases except the special case of null mutant activity in Fig. 6.3a where mutant alleles are at high frequency. The linear increase in the proportion of interaction variance in Fig. 6.3c is simply a consequence of a linear increase in the ratio of the number of two locus interaction terms to number of additive terms; higher order interactions contribute little variance because genotypes with more than two mutant alleles are very rare.

6.3.2. Branched Pathway

Variation in the Reference Branch. For the case of two variable loci in the reference branch, it can be shown from (6.6) that the flux in the reference branch has the same form as (6.11) except that the term K is replaced with,

$$K = b' + 1/(1/aK_{A,S} + 1/c), \tag{6.14}$$

where b' is the sum of reciprocals of enzyme activities in the reference branch excluding the activities of the two variable loci as in (6.4). Thus, the properties described for the linear chain are the same as described for this case. As $K \to \infty$, the control of the reference flux shifts to the enzymes in the non-variable segments of the system, and the proportion of interaction variance decreases.

With more than two varying loci in the reference branch, the interaction component changes with increasing number of segregating loci as in Fig. 6.3 (results not shown), i.e. V_I/V_T can have a maximum greater than that obtained when all loci are segregating.

Two Locus Interactions Between Loci in Different Branches. With two segregating loci, at least one of which is either in the common or competing branches, the

behaviour is rather different from the case of both loci in the reference branch. A change in activity of a locus in either the common or competing branches can lead in theory to a reversal in the direction of fluxes (see Kacser (1983) for an analysis of the behaviour of branched pathways in terms of control analysis). The values of a_1/ε and a_2/ε in (6.13) are no longer constrained by $1 < a_1/\varepsilon, a_2/\varepsilon < \infty$. The ratio of additive to interaction variances is a much more complex function and involves the concentrations of the externally controlled substances as well as equilibrium constants which are involved in determining the reversibility of the fluxes. This is illustrated in Fig. 6.4 which shows (a) the value of a_1/ε, and (b) the response of the reference flux as X_B is modulated. The activities of the wild type alleles at the two loci are equivalent and the mutant alleles are both nulls.

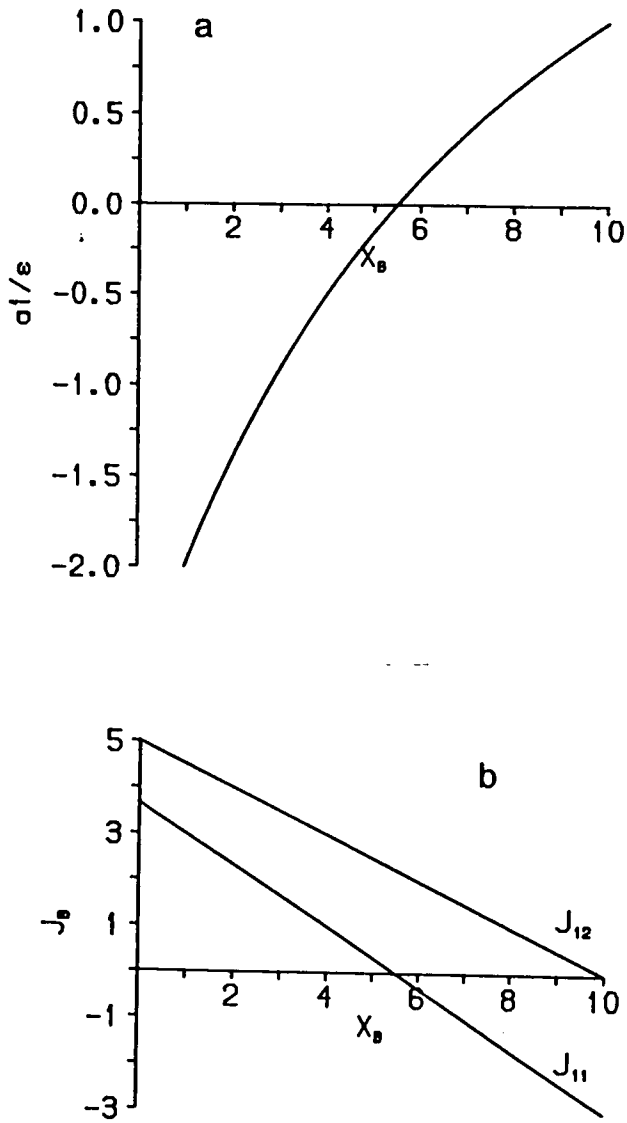


Fig. 6.4. The effect of asymmetry in the branched system. There are three enzymes in the system, one in each branch, and the wild type activity of each is one. In the "B" and "C" branches a mutant allele of zero activity segregates in the population at frequency 1/2. The value of X_A is 10 and X_C is 1. Fig. 6.4a shows the effect on the value of $a_1/\varepsilon = (J_{11} - J_{21})/(J_{11} - J_{21} - J_{12} + J_{22})$ of modulating X_B. Fig. 6.4b shows the effect on the value of the reference flux. J_{11} is the value of the double wild type flux and J_{12} is wild type for the "B" enzyme and mutant for the "C" enzyme.

As the value of the flux for the double wild type, J_{11}, approaches zero, the value of a_1/ε also approaches zero. Near this point, a change in direction of the flux occurs, and (6.13) indicates that the proportion of interaction variance can be substantial as $q_k \rightarrow 1$ (mutant allele rare), which contrasts with the result for variation in the reference branch only. Kacser (1983) showed that a change in direction of the flux can occur for a change in enzyme activity with a high control coefficient. The dependence of the proportion of interaction variance on the X's in the branched system contrasts with its independence in the linear pathway of X_A and X_B.

Multiple Locus Interactions in the Branched System. The effect on the interaction component of variance of many loci simultaneously segregating in each of the three branches is illustrated in Fig. 6.5. In this system, the thermodynamic pressures (i.e concentration differences in relation to equilibrium constants) for the reactions $X_A \rightarrow X_B$ and $X_A \rightarrow X_C$ are large relative to the thermodynamic pressure between the outputs. The Figure compares the effects of equivalent loci in a 'symmetrical' system, $X_B = X_C$ and in two 'asymmetrical' systems where $X_B >> X_C$ or $X_C >> X_B$, and can be compared to the interaction variance generated by a linear pathway (Fig. 6.3). The graphs have uneven shapes because the segregating loci are introduced into the three branches successively. Fig. 6.5 suggests that the effect of asymmetry is rather small because the three types of system give similar proportions of interaction variance. Most of the interaction in this case occurs between loci in the common ("A") branch and the reference ("B") branches because most of the increase in interaction occurs with introduction of segregating loci in these branches; the introduction of segregating loci in the competing branch does not lead to substantial increases in the proportion of interaction variance.

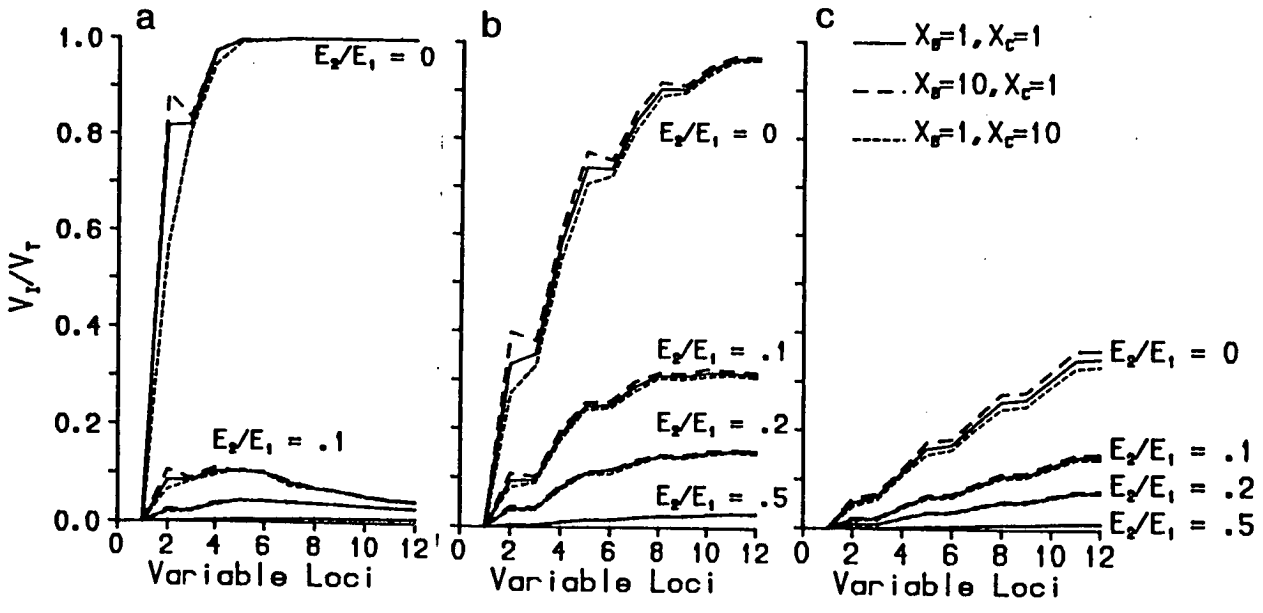


Fig. 6.5. The proportion of interaction variance in the reference flux of the branched pathway with variation simultaneously in three branches. Each branch has four equivalent enzymes. Each step on the horizontal axis corresponds to the introduction of one further variable locus into one branch and this occurs in the order A, B, C, A, and so on. The graphs relate various values of mutant to wild type activity ratio, E_2/E_1, and three types of system with varying degrees of asymmetry with respect to their thermodynamic properties. (a) q = 0.1; (b) q = 0.5; (c) q = 0.9.

6.3.3. Dominance Variance

We have been concerned primarily with the interaction component of non-additive variance, but in a diploid population a dominance component is present because of interactions between alleles within loci. The analysis of the interaction component also becomes more complicated because of the presence of additive by dominance and dominance by dominance interactions. These are higher order than additive by additive or dominance interactions, they are functions of third and higher derivatives of the population mean with respect to gene frequency and contain an additional $q(1 - q)$ term for each derivative. They are therefore likely to be small. It is worthwhile exploring the consequences of dominance in metabolic pathways due to two alleles segregating at one locus, and to evaluate the dominance component of variance thus extending the treatment of Kacser and Burns (1981).

Assume that the activity of the enzyme in the heterozygote is the average of the two homozygotes, which holds for most cases (Kacser and Burns, 1981). Using similar terminology to the haploid case, let the fluxes of the homozygous wild type, homozygous mutant, and heterozygote be $J_1 = C_X/(1/E_1 + K)$, $J_2 = C_X/(1/E_2 + K)$, and $J_3 = C_X/(2/(E_1 + E_2) + K)$ respectively, where $C_X = X_A - X_B/K_{A,B}$ and K is the sum of reciprocals of the non-variable loci as in (6.4). With Hardy-Weinberg equilibrium, the population mean is

$$m = q^2 J_1 + 2q(1 - q)J_3 + (1 - q)^2 J_2. \qquad (6.15)$$

Expressions for the additive and dominance variances can be obtained by Kojima's (1959) method. The ratio of additive to dominance variance is,

$$V_A/V_D = 2[q(1 - q)(\partial m/\partial q)^2]/[q^2(1 - q)^2(\partial^2 m/\partial q^2)^2] \qquad (6.16)$$

$$= 2[q + (J_3 - J_2)/(J_1 - 2J_3 + J_2)]^2/[q(1 - q)].$$

Substituting expressions for the fluxes of the three genotypes for the case of the linear chain gives,

$$V_A/V_D = 2[q - (1 + 1/(E_1K))/(1 - E_2/E_1)]^2/[q(1 - q)]. \quad (6.17)$$

Thus, the ratio of variances is a function of gene frequency, q, E_1K and E_2/E_1, and has the following properties.

(i) As the ratio of mutant to wild type activity approaches unity ($E_2/E_1 \rightarrow 1$), the variance becomes mainly additive.

(ii) As $E_1K \rightarrow 0$ the variance becomes mainly additive. With the non-variable enzymes condensed to a single step, E_1K can be rewritten as E_1/E_G with E_G the 'group enzyme activity'. As $E_1/E_G \rightarrow 0$, most of the control of the flux shifts to the variable locus implying a linear change of flux with change of E_1. The dominance deviation is therefore small.

(iii) If the heterozygote flux is nearly intermediate, the dominance variance is greatest relative to the additive at intermediate gene frequency ($q \rightarrow 1/2$).

(iv) As $E_2/E_1 \rightarrow 0$, the dominance variance is greatest relative to the additive when the mutant allele is rare ($q \rightarrow 1$). Note, this is the opposite of the condition for maximum additive by additive interaction variance in the haploid case.

These properties are illustrated in Fig. 6.6 which shows the proportion of dominance variance of flux, V_D/V_T, for three gene frequencies of the wild type allele and a range of values of E_2/E_1. For example, with five equivalent enzymes in the pathway ($E_1K = 4$) and the mutant allele at frequency 0.1 ($q = 0.9$), the proportion of dominance variance is almost 26% for a null mutant allele, but only 2% for $E_2/E_1 = 0.5$.

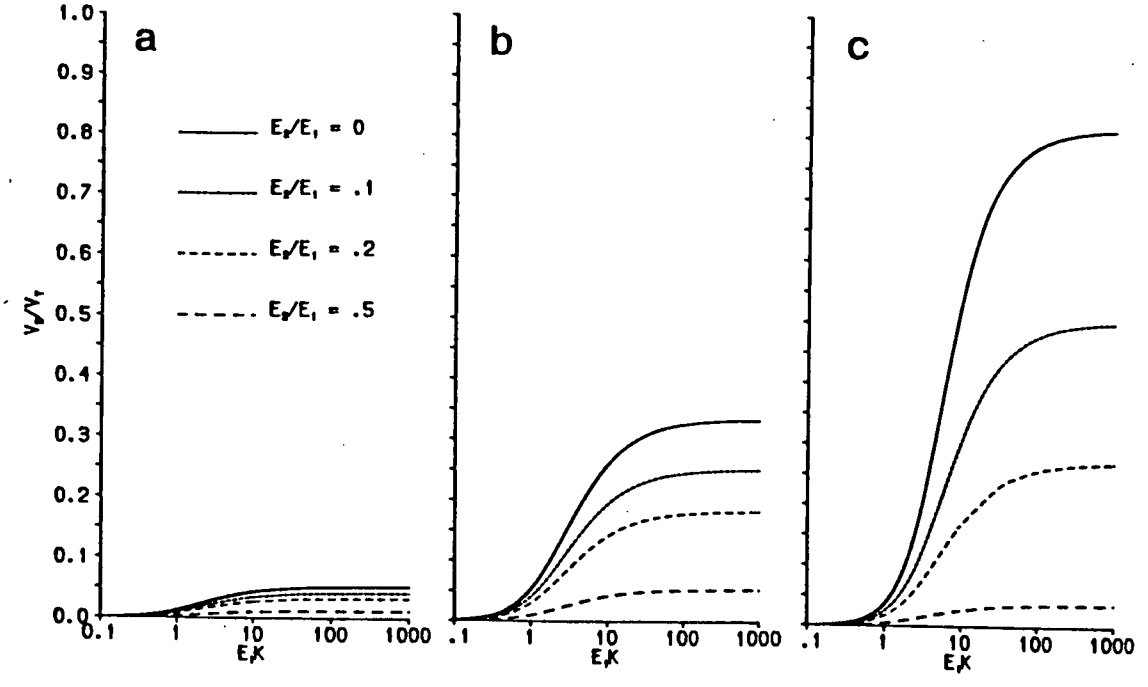


Fig. 6.6. The proportion of dominance variance of flux in a linear chain (see Fig. 6.1). The horizontal axis is the product of wild type activity at the segregating locus and the sum of reciprocals of activities of the rest of the loci in the chain. Each curve relates the variance component for different values of mutant to wild type activity ratio, E_2/E_1. (a) q = 0.1; (b) q = 0.5; (c) q = 0.9.

6.4. Discussion

6.4.1. Non-additive variance of flux

Allelic variation of enzyme activity generates both interaction and dominance variance of metabolic flux, but these are present in different circumstances. When the ratio of mutant to wild type activity is small, the proportion of interaction variance is maximal as the mutant alleles become common. The proportion of dominance variance in a diploid population is at a maximum as the mutant allele becomes rare. With a small difference between mutant and wild type, both non-additive components are maximal at intermediate gene frequencies.

There are implications of the above for responses to directional selection of metabolic flux. Assume that there is a range of allelic effects segregating at different loci controlling the flux in a pathway. Selection to increase the flux in the pathway will tend to fix the alleles of largest effect quickest. Their fixation leads to a reduction in their control of the flux because there is an inverse relationship between control of flux and activity at a locus. The loci at which alleles of smaller effect segregate must as a consequence increase their share of flux control, so the effect of allelic variation at these loci becomes larger (Kacser and Burns, 1981). This is similar to the epistatic effect described by Dykhuizen, Dean and Hartl (1987). Thus the epistasis allows new variation to be revealed as upward selection progresses, but a reverse attenuation effect occurs for selection of flux in the downward direction. In a haploid population, an asymmetrical response pattern is therefore expected. For a diploid population, directional dominance of alleles controlling flux is present. This causes an asymmetry in the opposite direction because the effects of mutant alleles are hidden in heterozygotes, so responses therefore show less asymmetry. The presence of non-segregating loci also affects dominance and additive by additive variances in different ways. As control of flux shifts to non-variable loci ($K \rightarrow \infty$), the proportion of interaction variance becomes small, but the proportion of dominance variance can become large as $E_2/E_1 \rightarrow 0$.

Where null alleles contribute to enzyme activity variation, the interaction component increases monotonically with increasing number of varying loci.

With mutant alleles of other than null activity, increasing the amount of allelic variation can actually decrease the proportion of interaction variance. This somewhat surprising result can be understood in terms of the distribution of flux control coefficients in the pathway. The introduction of further mutant alleles while producing an increase in the number of interactions between loci leads on average to a more linear change in flux from allelic substitution at any one locus. This linear change is less affected by allelic variation in enzyme activity at other loci.

The behaviour of the linear chain also applies to the branched system for alleles affecting enzyme activities within the branch in which the flux is measured (reference branch). The interaction component depends on the activity differences at the loci contributing to the flux variation, and also on the activities of the enzymes elsewhere in the system including the other branches, and decreases as their controlling influence on the flux increases.

The case of enzyme activity variation in different branches is more complex. In a branched pathway, changes in enzyme activity can, in theory, cause a change in the direction of flow, but since this situation is rather unlikely in nature, cases of large thermodynamic pressure from input (X_A) to outputs (X_B and X_C) are most relevant. In the branched pathway, the interaction component depends on the "externally determined" equilibrium constants and on the source (X_A) and sink (X_B and X_C) concentrations. It is therefore more difficult to generalize about epistasis in such a system. However, when the thermodynamic pressure from inputs to outputs is high relative to the pressure between the outputs, most of the interaction occurs between loci in the reference and the common branch. Presumably such loci contribute most substantially to the total variance, the loci in the competing branch being more kinetically distant (Kacser, 1983). A possible exception can occur when the reference flux is a small 'leak' relative to the competing flux. In this case, variation in the common and competing branches is likely to contribute substantially to the total genotypic variance in the reference flux and also to the interaction component should allelic activity differences be large. This "branch point effect" has been investigated previously by Kacser (1983) and by Laporte, Walsh and Koshland (1984).

6.4.2. *In vitro* enzyme activity differences

In this analysis, the specific activity of an enzyme is assumed to be proportional to the product of its turnover number and concentration. As well as changes in the structure of the enzyme molecule (which can affect turnover number and stability and hence its concentration), there are many classes of mutation capable of affecting both parameters. Such mutations include gene duplications and deletions, mutations at promoters, those affecting the stability and translation rate of mRNA, and mutations affecting the concentration and activity of 'effector' molecules. Thus, enzyme activity is itself a polygenic character for which there is evidence from studies of enzyme activity variation in natural populations of Drosophila (Laurie-Ahlberg *et al*, 1980; Laurie-Ahlberg, 1982; Graf and Ayala, 1986), and in mice (Paigen, 1979). Data on enzyme activity variation in natural and artificial populations falls into four main classes. (i) *In vitro* determination of activities of allozymes. (ii) *In vitro* activity differences between Drosophila chromosome substitution lines (e.g. Laurie-Ahlberg *et al*, 1980; Laurie-Ahlberg, 1982; Graf and Ayala, 1986; Miyashita and Laurie-Ahlberg, 1986). (iii) *In vitro* enzyme activity differences between inbred strains of mice. The measured enzyme activities vary widely between lines and the differences frequently depend on the tissue in which the activity is measured (e.g. Bulfield, Moore and Kacser, 1978; Johnson and Hong, 1986; Johnson, Hong and Knights, 1986). (iv) *In vitro* activity differences between artificially selected lines. Various enzymes activities thought likely to contribute to characters of economic importance in a number of species have been measured, including chickens (Bannister *et al*, 1984; Whitehead *et al*, 1984; Asante and Bulfield, 1988), mice (Asante, 1988), and pigs (Standal and Vangen, 1980). In the above studies, the largest range of enzyme activities has been detected between mouse inbred strains. In these and other studies, however, small and non-significant differences in activity were most common (i.e. less than a factor of four), with an increasingly small proportion of more extreme activity differences. The large activity differences measured between inbred mouse lines may reflect their polyphyletic origins (Bonhomme *et al*., 1987).

It is not possible to relate these observations directly to the biochemical models investigated here. These biochemical models are very simple approximations of *in vivo* pathways, a consequence of the need to incorporate a genetic model of enzyme activity variation segregating in a population.

However, the models show that small changes in enzyme activity (as observed in the *in vitro* studies) generate very little additive by additive interaction and dominance variance of flux as a proportion of the total genetic variance. Mutant alleles of low activity and at low frequencies contribute mainly dominance variance, but little interaction variance. This prediction of lack of interaction contrasts with the highly interactive nature of metabolic pathways within which enzymes are embedded. Conversely, the observation of substantial additivity for variation in a character does not necessarily imply that the components contributing to the variation are acting independently.

Chapter 7
Dominance in Metabolic Systems

7.1. Introduction

The characters which geneticists measure must depend on physiological and metabolic processes which occur within the organism. Genetic variation results in variation of the gene products of which many are catalytic proteins, i.e. enzymes. The metabolic map shows us how the enzymes are connected to one another by the metabolites they share. Other interactions are from metabolites that act as effectors, positive or negative, for particular enzymes. The map does not, however, give us any information on the rates at which substrates are converted from one to another. The measurement of these rates - metabolic fluxes - and of metabolite pool concentrations is the proper subject of 'quantitative metabolism'. The methodology of quantitative metabolism is to study the effects of varying the *parameters* of the system (e.g. enzyme activities) on *variables* which may be measured (i.e. fluxes and metabolite pool levels). Theoretical treatments of quantitative metabolism (Kacser and Burns, 1973; Heinrich and Rapoport, 1974; Kacser, 1983; Fell and Sauro, 1985; Hofmeyer, Kacser and Van Der Merwe, 1986) have provided expectations of the behaviour of living systems (e.g. Flint, Porteous and Kacser, 1980; Flint *et al.*, 1981; Groen *et al.*, 1982; Middleton and Kacser, 1983; Stuart *et al.*, 1986; Woodrow, 1986; Salter, Knowles and Pogson, 1986; Dykhuizen, Dean and Hartl, 1987).

Kacser and Burns (1981) applied the methods of quantitative metabolism to the question of the effects of finite changes in enzyme activity generated by allelic differences on dominance relationships of flux. It was shown that there is a non-linear relationship between flux or metabolite concentration and enzyme activity. The general expectation that 'null' mutants at enzyme loci are 'recessive' was explained in these terms without necessitating an evolutionary hypothesis of 'modifiers' first proposed by Fisher (1928) (see e.g. Middleton and Kacser, 1983; Dean, Dykhuizen and Hartl, 1986; Cornish-Bowden, 1987; Kacser, 1987). On the other hand, small differences in the enzyme parameters of mutant and wild type were shown to result in an intermediate heterozygote phenotype. Here, the Kacser and Burns treatment is extended to the problem of

pleiotropic effects of enzyme variation.

Because of the interconnected structure of the metabolic system, genetic variation at one locus in principle affects all the measurable characters and there is therefore a general expectation of pleiotropy. It does not, however, imply that any genetic variation affects all the characters in the same direction or to the same extent. The metabolic map shows that some characters are 'closer' to one another and others are more 'distant'. The purpose of this chapter is to model the effects of finite changes in genetically determined enzyme parameters on characters pleiotropically coupled in the metabolic system. In particular, expectations for the dominance relations will be derived. The results will be related to the observation that, in a vast majority of cases, the dominance relations of a pair of alleles are the same for all the pleiotropically affected characters which can be measured.

The dominance relationship of the three phenotypes in a diploid can be described by the Dominance Index, D, first defined by Wright (1934). Using similar terminology as in Ch. 6, the values of the three diploid phenotypes are defined as J_1, J_2 and J_3 for the wild type, mutant and heterozygote respectively. D is defined by

$$D = (J_1 - J_3)/(J_1 - J_2). \tag{7.1}$$

Although D can take any value, the following useful conditions apply for limiting cases: (1) $J_1 = J_3$, then $D = 0$, the mutation is fully 'recessive'; (2) $J_1 - J_3 = (J_1 - J_2)/2$, then $D = 1/2$, the mutation gives an exact intermediate heterozygote phenotype; (3) $J_3 = J_2$, then $D = 1$, the mutation is fully 'dominant'.

7.2. Biochemical models

The simple case of a branched pathway is investigated (Fig. 7.1).

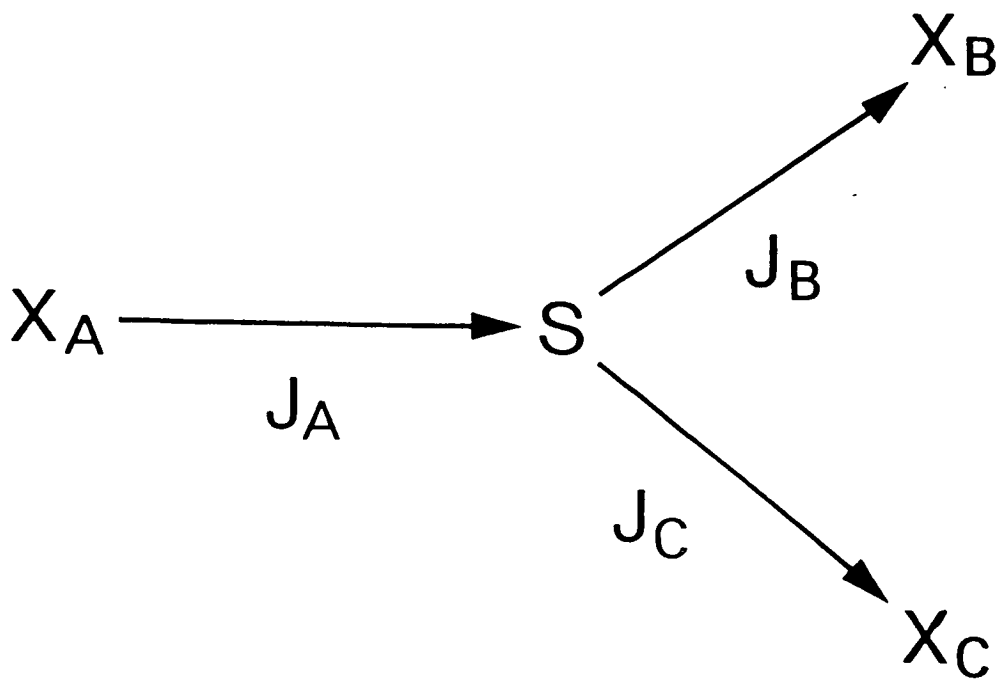


Fig. 7.1. The branched system. This consists of three linear branches. S is the product of the common branch "A" and is the substrate of the reference "B" and competing "C" branches.

This simple biochemical system shows pleiotropy, as two outputs are affected by changes in any enzyme activity. It is described in more detail in Ch. 6 (Fig. 6.2). Although even an unbranched pathway displays pleiotropy, insofar as enzyme variation can affect the intermediate pools differentially, a branched system with two outputs is a more general case. The system is assumed to be at steady state with the concentrations of input substance X_A and outputs X_B and X_C constant. At steady state the fluxes (J's) through each pathway are constrained by

$$J_A = J_B + J_C, \tag{7.2}$$

(cf. Ch. 6). Enzyme variation anywhere in the system simultaneously affects all the fluxes and metabolite concentrations though not necessarily to the same degree.

7.2.1. Unsaturated system

Fig. 7.1 represents the general case of a branched pathway. By making the simplifying assumption that all the steps are monomolecular, and that saturation of the enzymes is absent, it is possible to derive a system of linear equations for the fluxes and pools (see Ch. 6). As a further simplification the activities of the enzymes in each branch can be combined to give 'group enzyme activities' a, b and c. For the unsaturated system, there are therefore three linear rate equations for the three fluxes. Since the system is symmetrical, the flows in any one branch can go on in either direction as long as the mass conservation constraint (7.2) is met. In the following sections, the following directions are defined for positive fluxes: $X_A \rightarrow S$, $S \rightarrow X_B$, $S \rightarrow X_C$. By solving the simultaneous equations, expressions for each of the fluxes can be obtained in which the enzyme activities appear as parameters. Consider the two allele case: wild type, mutant and heterozygote for which three different values of one of the enzyme activities apply. By inserting these three different enzyme values in each of the equations for J_A, J_B and J_C, the phenotypic values for the fluxes are obtained. Finally, insertion of these in to the expression for the Dominance Index (7.1) allows us to compare how changes at one locus affects the dominance relations for the pleiotropically related fluxes.

In the Appendix a general proof is given that for a system of unsaturated enzymes, of any complexity, the Dominance Index with respect to any one locus is exactly the same for all characters (fluxes or metabolite pools) affected by the allelic substitutions. Although the Dominance Indices are identical, the measured differences in the characters may be very different. If, for example, we find 'recessivity' in one character, say a flux, ($D \rightarrow$ O), then we shall find the same in another pleiotropically related character, say a pool. The underlying feature which generates these identities is that in a system with linear equations for each step, all the fluxes are linearly related to all the pools. This is illustrated in Fig. 7.2 which is shown as a 'reflection diagram' (see Burns and Kacser, 1977). Such a diagram shows the effects of the independent variable (group enzyme activity $1/a$ in this case) on a dependent variable (S in this case) which, in turn, affects further dependent variables (the three fluxes). The decomposition into the functional components aids the understanding of the system.

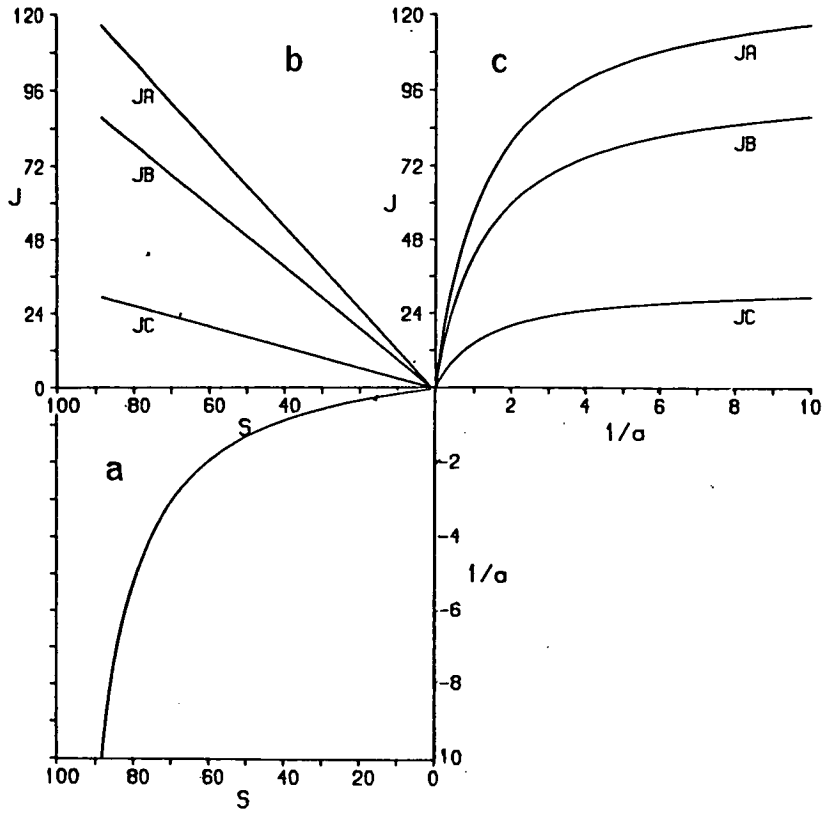


Fig. 7.2. Flux responses to changes in the enzyme activity of the common branch in a system with first order enzymes (all enzymes unsaturated). The effects are shown as a 'reflection diagram' which decomposes the effect into component functions of the final function. Fig. 7.2a shows how variation in the parameter $1/a$ (common branch group enzyme activity) changes the steady state value of the variable S (the concentration of the branch point substrate). In Fig. 7.2b, changes in S are now reflected in this panel which shows how changes in the variable affect the three fluxes dependent on it. This therefore shows how S is allocated from J_A to the two output fluxes J_B and J_C. In Fig. 7.2c, the flux changes are reflected back to $1/a$ and show the resultant net effect of changes in $1/a$ on the three fluxes. It will be noted that Fig. 7.2b shows linear relationships of S on J's. This implies that the Dominance Index (resulting from any three values of a) will be identical when measured in any of the three fluxes or in S. The relationship of the J's on $1/a$ in Fig. 7.2c are simple hyperbolic functions. Any three values of $1/a$ will give identical flux proportions in all three fluxes. It will also be noted that the fluxes are constrained by $J_A = J_B + J_C$.

7.2.2. The branched system with saturation

Saturation or feedback inhibition are two features of metabolic systems which introduce non-linearity of rates of individual steps with respect to their substrates. In the 'linear' system, the rate of each step is linear with respect to the substrate concentrations, but the fluxes are non-linear functions of the enzyme activities. Consider the case of some degree of saturation in one of the branches (Fig. 7.3).

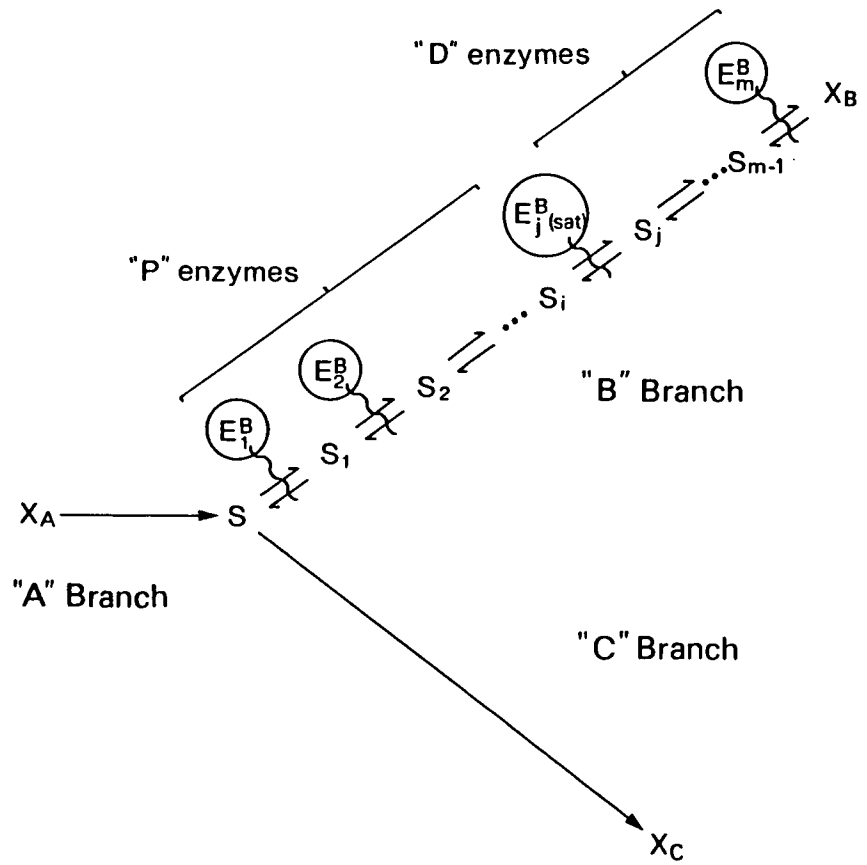


Fig. 7.3. The saturable branched system. As before there are three branches labelled "A", "B" and "C". The "A" and "C" branches are identical in structure to the branches as specified for Fig. 6.2 and are assumed to be unsaturated. The "B" branch consists of three parts. The linear "P" section is proximal to the saturable step $S_i \rightarrow S_j$ catalysed by enzyme E_j. Distal to this step is the "D" portion of unsaturated enzymes. The nomenclature is the same as described in the legend to Fig. 6.2.

The "A" and "C" branches are identical in structure to the non-saturable system. The saturable branch can be divided into three parts. (a) The "P" section of non-saturable enzymes *proximal* to E_j. (b) The saturable enzyme E_j. E_j is saturable by either its product S_j or its substrate S_i or both, depending on the values of the Michaelis constants, M_j and M_i, respectively. (c) The "D" section of non-saturable enzymes *distal* to E_j. At steady state, there are four equations for the fluxes in the linear segments of the system:

$$J_A = (X_A - S/K_{A,S})/a, \tag{7.3}$$

$$J_C = (S - X_C/K_{S,C})/c, \tag{7.4}$$

$$J_B = (S - S_i/K_P)/p, \tag{7.5}$$

$$J_B = (S_j - X_B/K_D)/d. \tag{7.6}$$

where $p = \Sigma 1/E_i^P$, and $d = \Sigma 1/E_i^D$ represent the (linear) 'group enzyme activities' in the "P" and "D" sections respectively, and K_P and K_D are the equilibrium constants for the reactions $S \rightarrow S_i$ and $S_j \rightarrow X_B$ respectively. The flux for the step with saturation containing the relevant Michaelis constants is given by (6.1). Elimination of the variable S terms and substitution into (7.2) gives a quadratic formula for J_B,

$$\begin{aligned} &J_B^2[dM_i/M_j - K_P(p + 1/(1/(aK_{A,S}) + 1/c))] \\ &+ J_B[dV_j/K_{i,j} + X_B M_i/(M_j K_D) + V_j K_P(p + 1/(1/(aK_{A,S}) + 1/c)) \\ &\quad + M_i + K_P(X_A/a + X_C/(cK_{A,C}))/(1/(aK_{A,S}) + 1/c)] + \\ &X_B V_j/(K_{i,j} K_D) - V_j K_P(X_A/a + X_C/(cK_{A,C}))/(1/(aK_{A,S}) + 1/c) = 0. \end{aligned} \tag{7.7}$$

For an asymptotically stable steady state this has only one positive real root.

The expression for J_C is obtained by eliminating S from equations (7.2), (7.3) and (7.4) to give,

$$J_C = (X_A - X_C/K_{A,C} - aJ_B)/(a + c/K_{A,S}), \tag{7.8}$$

where $K_{A,C}$ is the equilibrium constant for the reaction $X_A \rightarrow X_C$. The root of equation (7.7) (if known) can then replace J_B in (7.8) giving an expression for J_C in terms of parameters only.

The effects of saturation on the dominance relations of the outputs, J_B and J_C, are examined for cases of 'high' and 'low' saturation. The degree of saturation may be described quantitatively by a saturation index, SAT

$$SAT = (S_i/M_i + S_j/M_j)/(1 + S_i/M_i + S_j/M_j). \tag{7.9}$$

SAT can take values as follows: (i) $S_i/M_i << 1$ and $S_j/M_j << 1$ ('Low' saturation, $SAT \rightarrow 0$.) (ii) $S_i/M_i >> 1$ and/or $S_j/M_j >> 1$ ('High' saturation, $SAT \rightarrow 1$.) Differences in saturation were investigated by modifying a parameter, X_A, which in turn affects the amount of saturation. As X_A increases, the saturation of E_j increases. This is illustrated in Fig. 7.4. (An alternative method would have been to compare a series of enzymes with decreasing values of Michaelis constant, M).

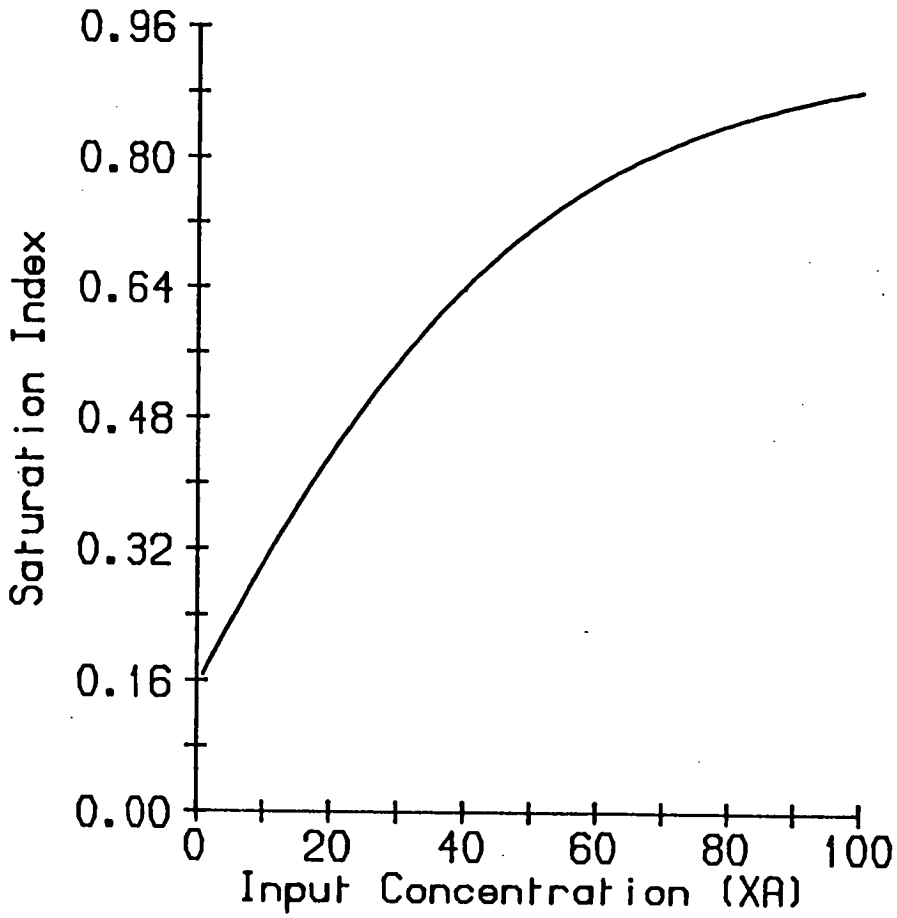


Fig. 7.4. Saturation Response of E_j to X_A. The following parameters were used: X_A varied from 1 to 100; X_B= 1; X_C = 1; a = 0.5; p = 0.01; d = 0.01; c = 1; V_j = 100; M_i = 5; M_j = 1000. All equilibrium constants were set to unity. These values were used to compute the fluxes and the pools, S_i and S_j, which were then inserted into (7.9).

Enzyme Variation in the Common Branch. Fig. 7.5 shows Dominance Indices of J_B and J_C for enzyme variation in the common branch as X_A is modulated and hence as the degree of saturation changes. With high levels of saturation, the Dominance Indices can be quite different with, in this case, J_B more 'recessive' (smaller D). At low saturation, as the analysis of the unsaturated case predicts, the Indices tend to equality. The differences in dominance which can occur are explained in terms of the non-linearity of the fluxes to substrate concentrations now present in the system. Since the competing pathway is a chain of non-saturable enzymes, J_C responds linearly to changes in the common substrate, S, but J_B responds non-linearly due to the damping effect of the saturable enzyme. This is illustrated in Fig. 7.6 which is shown as a 'reflection diagram' for one value of X_A giving high saturation.

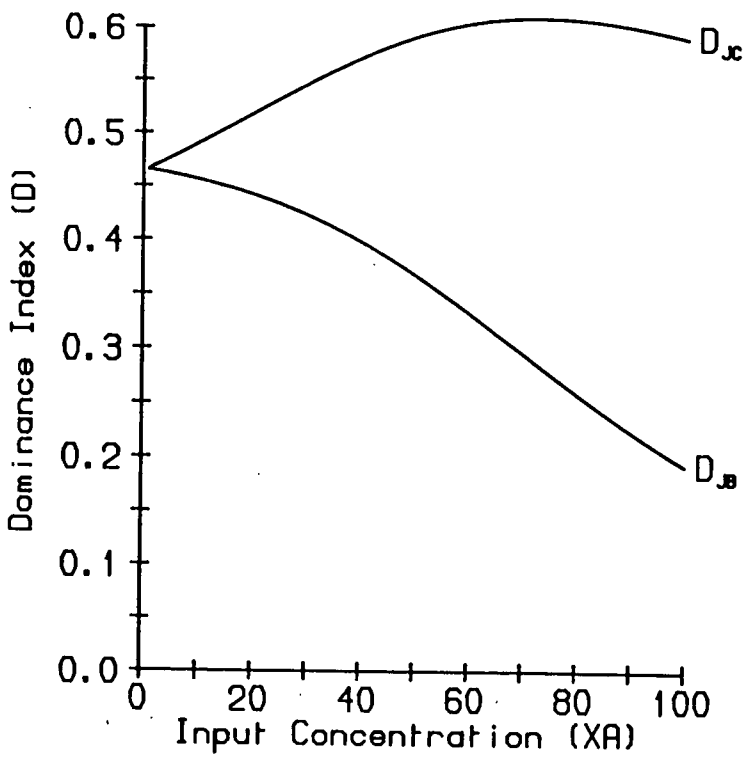


Fig. 7.5. Dominance Indices as a function of saturation by varying X_A. The Dominance Indices of the two output fluxes with respect to enzyme variation occurring in the common ("A") branch. The parameters were identical to those used in generating Fig. 7.4, except for $1/a$ which had values of mutant activity = 0, heterozygote activity = 1 and the wild type activity = 2.

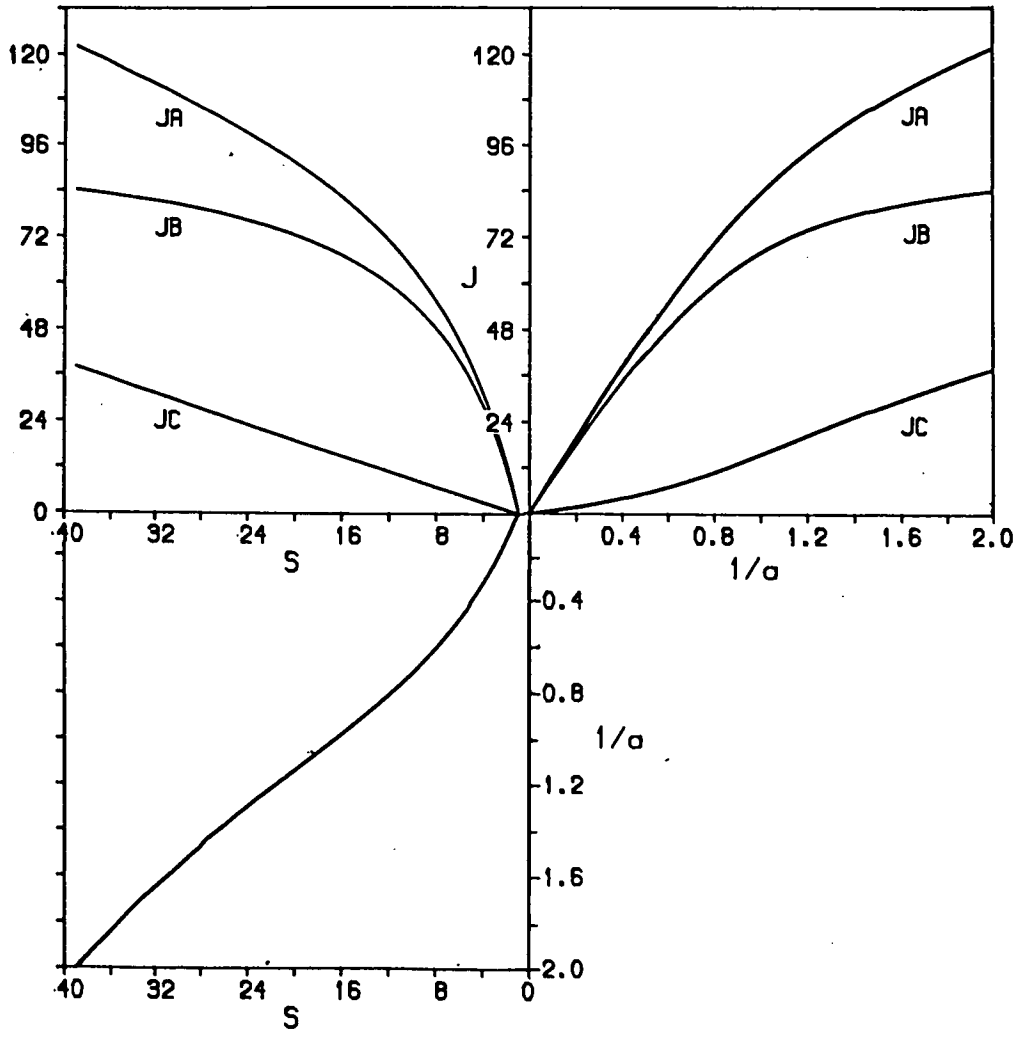


Fig. 7.6. Flux responses to changes in the enzyme activity of the common branch. Parameters as in Fig. 7.4 with X_A at 100 (high saturation) and $1/a$ varying from zero to 2. As explained in the legend to Fig. 7.2, the effects are shown as a 'reflection diagram'. The competing branch of first-order enzymes gives a linear flux response to S. The reference branch, however, gives a damped response due to the presence of the saturable enzyme. This damping effect leads to the tendency for the reference flux to give a 'more recessive' phenotype than the competing flux (see Fig. 7.5) and hence to different Dominance Indices.

Fig. 7.6a shows the effect of varying $1/a$ on the common substrate S. Unlike the response in non-saturated systems, where the relationship was hyperbolic (see Fig. 7.2), S shows an early 'accelerating' portion before approaching a plateau at high values of $1/a$. Fig. 7.6b shows the effect of such changes in S on its differential allocation to the two outputs. J_C responds linearly to S (cf. section on non-saturated systems), while J_B shows the effects of increasing saturation. Finally, Fig. 7.6c shows the net effect of changes in $1/a$ on the fluxes. Thus, no matter which three values of S result from the three enzyme activities, the J_B phenotype will appear 'more recessive' than J_C. It is notable that in this case the flux in the non-saturable branch, J_C, can give a Dominance Index greater than 1/2, i.e. the mutant can therefore tend to be 'dominant' over the wild type. This result is due to the increasing slope of the J_C flux at low values of $1/a$ (Fig. 7.6c) in contrast to the monotonically declining change in J_B.

Enzyme Variation in the Non-saturable Branch. Fig. 7.7 illustrates the Dominance Indices of the two outputs for varying values of X_A where enzyme variation occurs in the competing branch. Here, the opposite result is observed from the effect of variation in the common branch. The flux in the non-saturable pathway, J_C, is in this case a more recessive phenotype than that measured in the saturable branch. This is also explainable in terms of the response of the fluxes to changes in the common substrate, S, illustrated in Fig. 7.8. As in the case of variation in the common branch, J_B varies non-linearly with S, higher values being damped due to the effect of saturation (see Fig. 7.6b). The flux in the common branch, J_A, is, however linear in S. Since $J_C = J_A - J_B$, J_C varies non-linearly with S, changes in which have been induced by enzyme variation in this pathway.

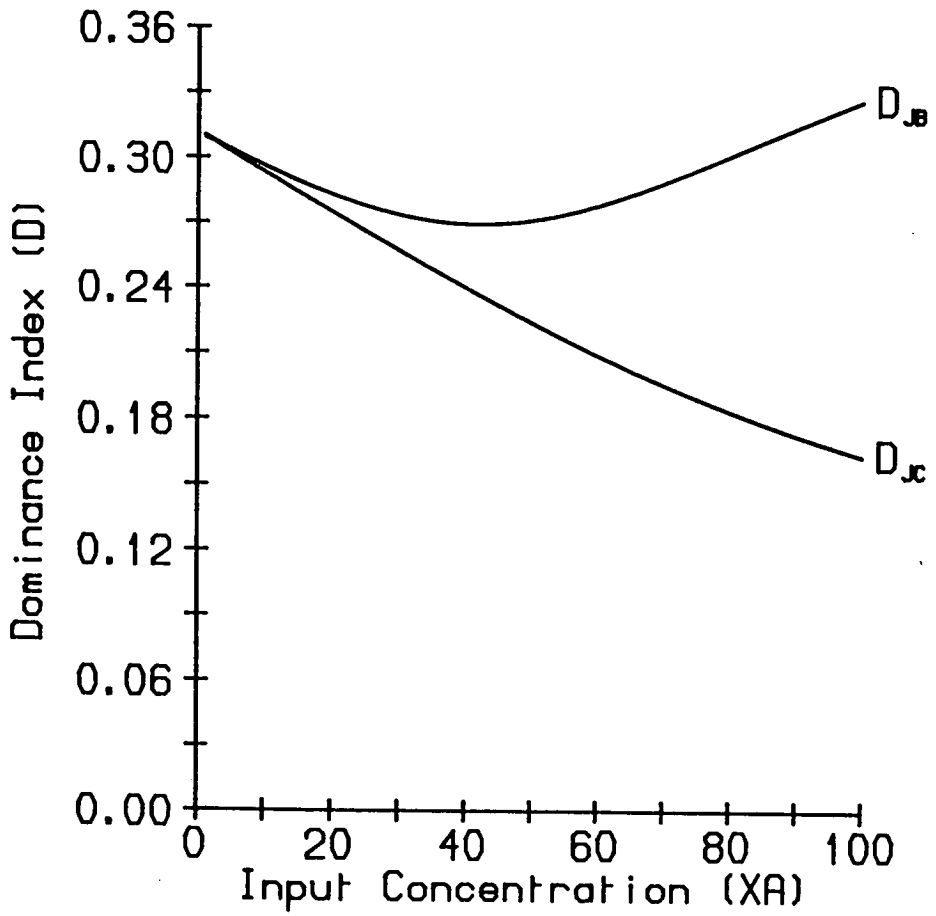


Fig. 7.7. The Dominance Indices of the two output fluxes with respect to enzyme variation occurring in the non-saturable competing branch. Saturation is varied in the system by modulating the input X_A. The mutant enzyme activity is zero and the heterozygote activity is half wild type. The following parameters were used to generate the curves: X_B = 1; X_C = 1; a = 1; p = 0.01; d = 0.01; c = 0.1 (wild type); V_j= 50; M_i = 5; M_j = 1000. All equilibrium constants were set to unity.

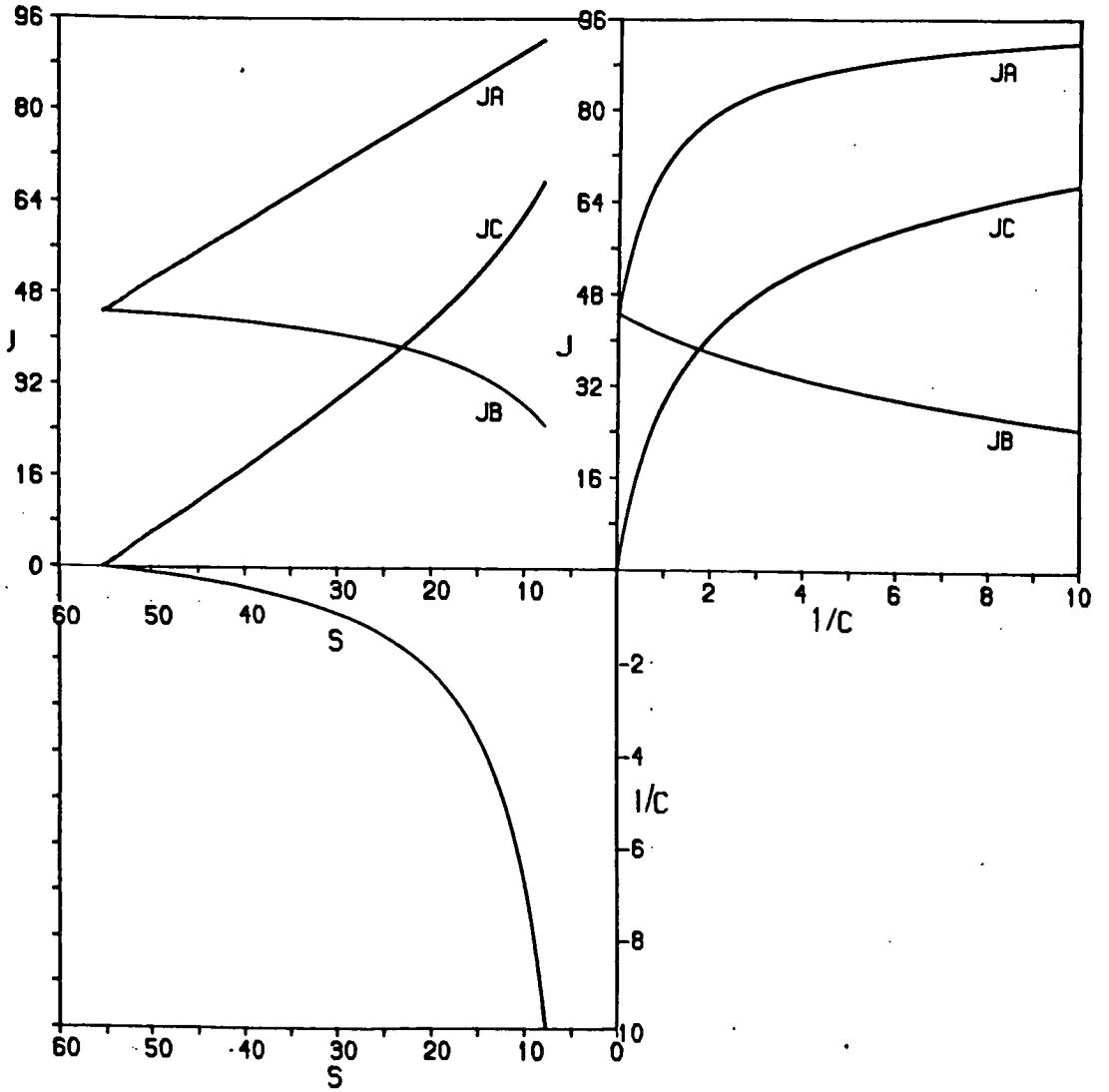


Fig. 7.8. Flux response to changes in enzyme activity in the competing "C" branch in the saturable system. This is a 'reflection diagram' (see Fig. 7.2). Parameters as in Fig. 7.7 with X_A = 100 and $1/c$ varying from zero to 10. The common flux gives a linear response to changes in S, and reference flux is damped due to the presence of saturation. The net result of these changes is the tendency for J_C to give a more recessive phenotype than J_B.

In this case the numerical studies indicated, somewhat surprisingly, that the Dominance Indices measured in each character were identical although there was considerable saturation in the branch. This is so irrespective of whether enzyme variation occurs before, after, or at the saturable step. The explanation lies in the linearity of the fluxes in the non-saturable branches (J_A and J_C) with changes in S caused by enzyme variation in the saturable branch. Since the fluxes are constrained by $J_B = J_A - J_C$, the reference flux must also be linear with changes in S caused by enzyme variation in its own pathway (Fig. 7.9). (Contrast Figs. 7.6b and 7.8b).

In the simulations it has been assumed that a mutation of the saturable enzyme affects only the V_{MAX}, and not the K_m's. This is an unnecessary restriction as changes in any enzyme parameter affects the phenotypes through changes in S. The behaviour of the indices is therefore the same as discussed above.

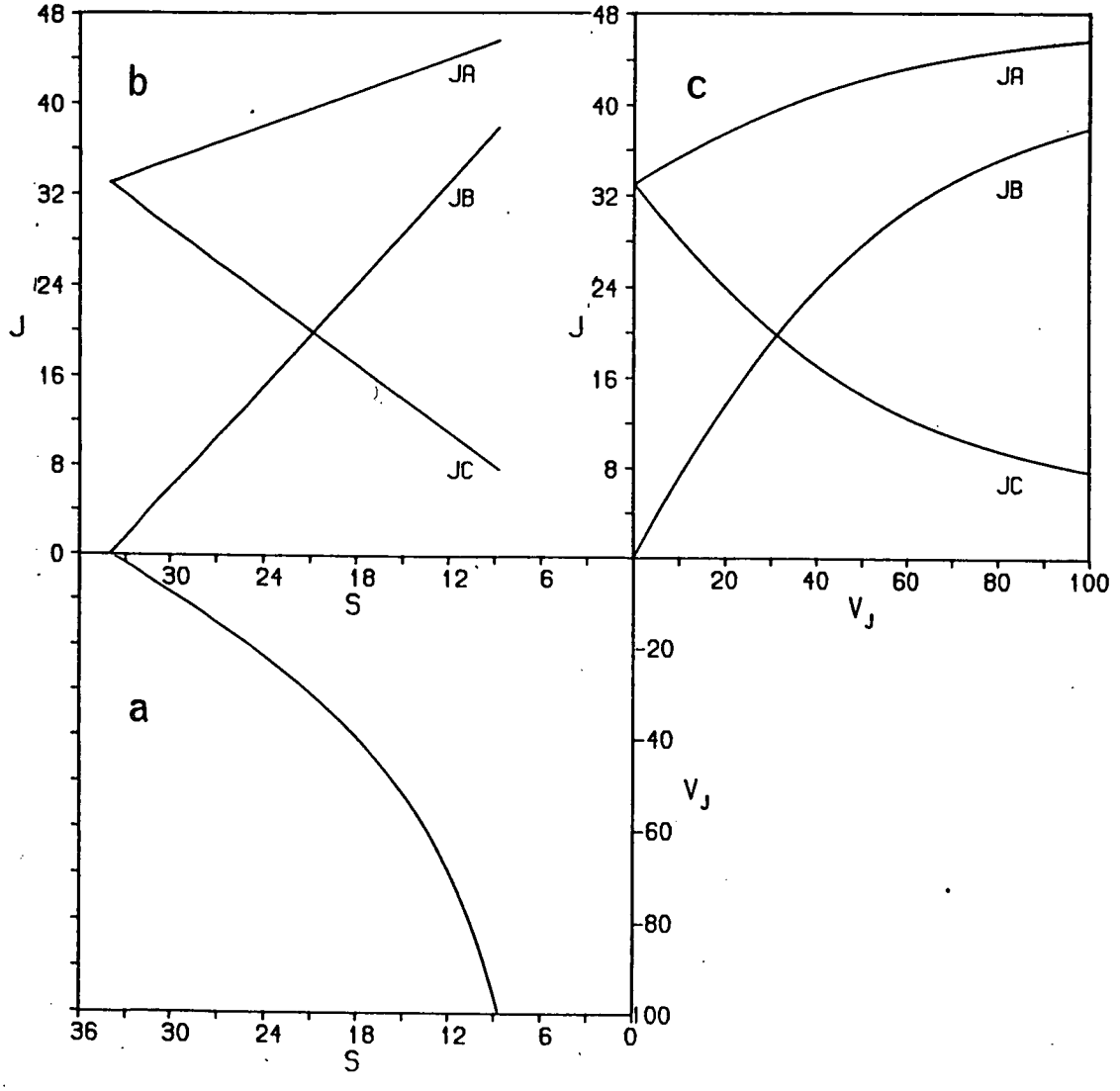


Fig. 7.9. Flux responses to change in enzyme activity in the saturable "B" branch. Parameters: X_A = 100; X_B = 1; X_C = 1; a = 2; p = 0.01; D = 0.01; M_i = 10; M_j = 1000; c = 1; and V_j varied from 0 to 100. All equilibrium constants were set to unity. Here the relationships of the fluxes to the branch point metabolite, S, are all linear, leading to identical dominance relations.

7.2.3. The Effect of Feedback Inhibition

Feedback inhibition is another mechanism which can cause non-linear relationships between pools and fluxes. The effect of a feedback inhibition loop (Fig. 7.10) on the dominance relations of the output fluxes has also been investigated.

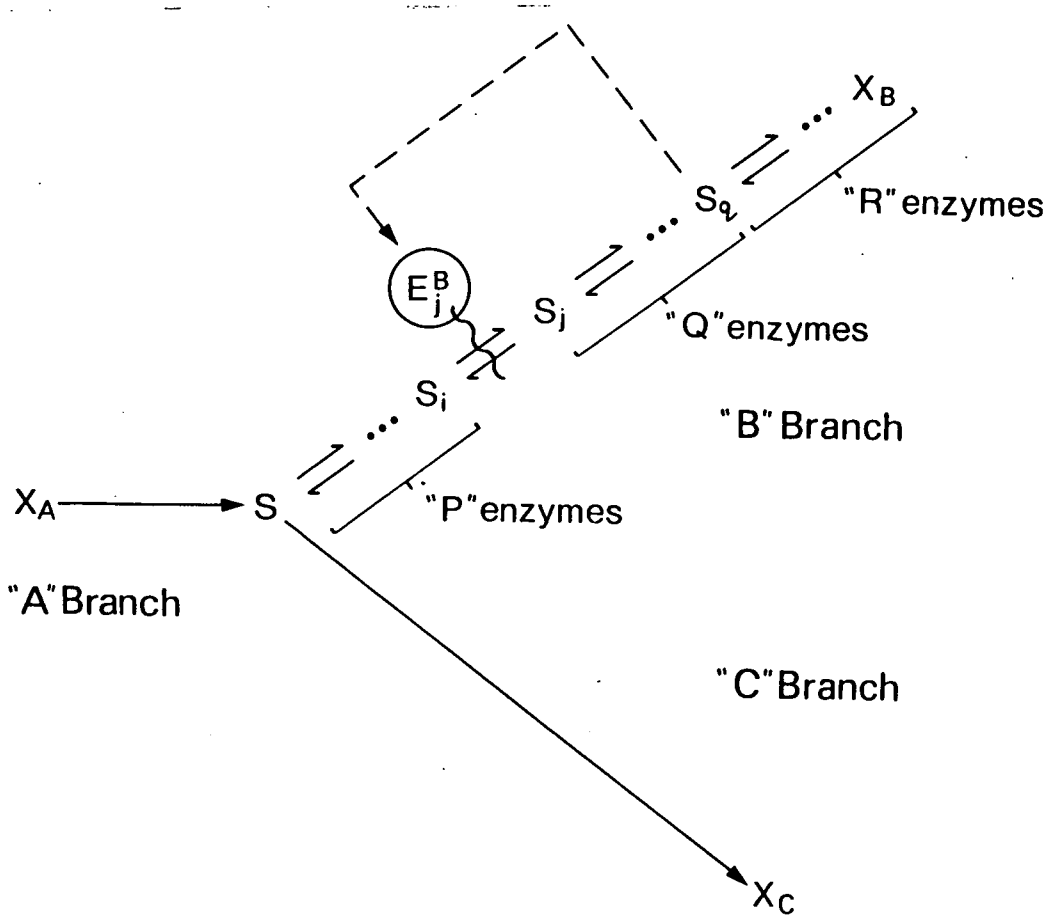


Fig. 7.10. The Branched System With Feedback Inhibition. The structure is similar to the previously defined systems (Figs. 7.1 and 7.3). The "A" and "C" branches are identical in structure to the linear system. The "B" branch consists of four parts. The "P" section is a chain of linear enzymes proximal to the step $S_i \rightarrow S_j$ catalysed by enzyme E_j. This enzyme is inhibited by S_q, a substrate further up the chain. Proximal to S_q is the "Q" section of linear enzymes and distal is the "R" section. The equilibrium constants for the reactions $S \rightarrow S_i$, $S_j \rightarrow S_q$, and $S_q \rightarrow X_B$ are termed K_P, K_Q and K_R respectively.

An expression for the reference flux with feedback inhibition in the reference branch is derived as follows. At steady state, the set of simultaneous equations defining the fluxes is (7.2), (7.3) and (7.4) together with the following equations for J_B,

$$J_B = (S - S_i/K_P)/p, \tag{7.10}$$

$$J_B = (S_j - S_q/K_Q)/q, \tag{7.11}$$

$$J_B = (S_q - X_B/K_R)/r, \tag{7.12}$$

where the equilibrium constant (K) terms are as defined in the legend to Fig. 7.10 and $p = \Sigma 1/E_i^P$, $q = \Sigma 1/E_i^Q$ and $r = \Sigma 1/E_i^R$ represent the 'group enzyme activities' of the "P", "Q" and "R" sections respectively. A simple expression for the rate of the reaction $S_i \rightarrow S_j$ catalysed by E_j is

$$J_B = (V_j/M_i)(S_i - S_j/K_{i,j}) \; / \; (1 + S_q/K_I) \tag{7.13}$$

(Cleland, 1963), where S_q is the concentration of the metabolite which acts as an allosteric inhibitor, and K_I is the inhibition constant. The solution to the set is a quadratic in J_B,

$$\begin{aligned} & J_B^2 r/K_I + \\ & J_B[1 + X_B/(K_R K_I) + V_j p K_P/M_i + V_j q/(M_i K_{i,j}) + \\ & \quad V_j r/(M_i K_{i,j} K_Q) + V_j K_P/(M_i(1/(aK_{A,S}) + 1/c))] + \\ & V_j X_B/(M_i K_{i,j} K_Q K_R) - \\ & \quad V_j K_P(X_A/a + X_C/(cK_{S,C}))/(M_i(1/(aK_{A,S}) + 1/c)) = 0. \end{aligned} \tag{7.14}$$

The flux through the competing pathway is given, as before, by (7.8).

The non-linearity introduced by feedback inhibition leads to qualitatively similar

behaviour to the system with saturation in one branch (results not shown). As feedback inhibition increases (measurable by an index analogous to the saturation index), differences in dominance measurable in the outputs can occur, of the same direction as with saturation. The arguments to explain these phenomena are identical to those used to explain the effects of saturation.

7.3. Discussion

As a limiting case, dominance in systems of monomolecular transformations with no saturation or feedback inhibition has been investigated. This is a reasonable approximation to systems when the metabolite concentrations are less than their respective Michaelis constants. The conclusions apply to any arbitrary network of any structural complexity. An important result is that this type of system has identical Dominance Indices for all fluxes and pool levels with respect to variation of any of the enzymes. The reason for this behaviour is that in such systems, the pools and fluxes are linearly related to one another. Real metabolic systems are not of course all monomolecular as many reactions involve splitting and combining substrates. For example, a bimolecular step in a divergent metabolic system could be the reaction $S_i \rightarrow S_j + S_k$, where S_j and S_k are the beginnings of two further pathways. This reaction is constrained by its stoichiometry and the rates of production of S_j and S_k are necessarily identical. If two outputs arise from such stoichiometrically constrained fluxes, no competition for a shared metabolite occurs, so no difference in dominance relationship will be observed, no matter what the saturation state or feedback conditions are.

For non-linear systems numerical simulations have been used to assess the effects on independently variable fluxes. It has been shown that saturation at one enzyme can lead to differences in the dominance relations of two outputs. The directions of the differences in Dominance Indices for enzyme variation at different parts of the system have been defined. Systems with saturation in both the competing branches have not been explicitly considered. The directions of any differences in dominance will depend on the relative saturation in each of the branches. The dominance relations in the presence of such non-linear enzymes are dependent on their Michaelis constants as well as on the concentrations of external substances which affect the degree of

saturation. Dominance and any possible differences are thus a function of the 'environment' as well as of the genes.

The effect of saturation on the Dominance Indices is essentially through the resulting non-linearity of fluxes to substrate concentration. The presence of a feedback inhibition loop in the system can also lead to differences in the dominance between the outputs. The mechanism is essentially the same as for the case of saturation, with feedback inhibition introducing non-linear relationships between fluxes and pools.

How do the above considerations relate to the dominance relations for pleiotropic effects of enzyme variation actually found *in vivo*? Human inborn errors of metabolism are frequently caused by null-mutations at enzyme loci (Harris, 1980). In most cases, heterozygotes are detectable if measurement of the enzyme activity is possible when it will show an activity which is usually the mean of wild type and mutant (Stanbury *et al*, 1983). In spite of this, the clinical phenotype of the heterozygote is 'recessive' and cannot be distinguished from the wild type. The heterozygote phenotype , however, invariably shows a small average difference from the wild type (e.g. Knox and Messinger, 1958; Bulfield and Kacser, 1974). Detection of heterozygotes therefore depends on the ability to measure small differences and on the environmental noise in the system. No case in the literature has been found where the heterozygote for a clinical phenotype is distinguishable from the wild type in the 'main effect' or any other manifestation of autosomal genes.

In Drosophila many null-mutants at enzyme loci are known. In these cases the pattern appears to be the same as that for human inborn errors of metabolism: recessive alleles are recessive for all their pleiotropic effects (Lindsley and Grell, 1968).

There are, however, some cases in the literature where mutants show different dominance relations for different characters. In mice, for example, homozygotes for W^f alleles at the W locus are extensively depigmented, and there is no obvious pattern to the depigmentation (Guenet *et al*, 1979). Heterozygotes resemble the wild type except, however, for the presence of white spots on the forehead and belly. In these areas, the mutant allele is therefore 'dominant' for hair colour, and in the rest of the animal, it is 'recessive'. The 'internal environment' of the gene in the different tissues must be different.

A further example of differences in dominance relations occurs in the halothane gene in pigs (reviewed by Webb, 1981). Homozygotes for the mutant allele are sensitive to halothane, are stress-susceptible, have improved meat colour and have improved performance for a number of economically important traits. Heterozygotes are not detectably different from the wild type for sensitivity to the anaesthetic, stress susceptibility, or meat colour, but are nearly intermediate for the other traits (e.g. growth rate and carcase quality). Apparently therefore, the allele is 'recessive' for some of the characters it affects, but is 'additive' for others. One less clearly described example of this type of phenomenon has been found in *Drosophila melanogaster*. The allele *scabrous-like* found in an abdominal bristle number selection line (Hollingdale, 1971) is a recessive semi-lethal, but has a substantial effect on bristles in the heterozygote. The biochemical bases of all these effects are of course obscure. On the whole, however, it would appear from the literature that most mutations, especially where the lesion is at an enzyme locus, have similar dominance relations for the characters they affect pleiotropically.

How relevant are the conclusions from the very simple models to the very much more complex *in vivo* metabolic system and to the 'characters' arising from its operation? The results show that four conditions must be *simultaneously* satisfied if substantial differences in dominance of pleiotropically related characters are to be observed.

i) Non-linearity of metabolites to fluxes. The fact that, in principle, all enzymes are non-linear converters and that feedbacks are frequent features of metabolism is not in itself a sufficient condition. Saturation must be high or the feedback function must be steep for significant deviation from linearity to occur. Evidence concerning *in vivo* saturation is very sparse, but the available data (e.g. Flint, Porteous and Kacser, 1980; Hess, 1973) suggest that most enzymes operate below or near their substrate Michaelis constants, although co-factors, such as NAD, appear to be present in saturating concentrations. Strong non-linearities may therefore be an exception rather than the rule.

ii) Non-linearity must occur in a branch other than that in which the allelic variation occurs.

iii) The fluxes must not be stoichiometrically constrained, as such fluxes do not give rise to differences in dominance.

iv) Finally, the variation must affect a step which is reasonably sensitive to changes in enzyme activity. This means that the heterozygote phenotype has to show a clear difference from the wild type (i.e. the mutant must not be effectively 'recessive'). If such complete recessivity of the flux through the enzyme obtains, it will necessarily imply that the pools (including the pool at the branch point) will show no variation in the heterozygote. Since any effect on other pleiotropic fluxes is only mediated via a change in the branch pool(s), such other fluxes will also show recessivity and no substantial difference in dominance can arise. In terms of the concepts of control analysis, the affected step should have a reasonably high control coefficient.

In vivo such steps with high coefficients are relatively rare (Kacser and Burns, 1973, 1981). In the simulation, the enzymes in the branches were 'condensed' to a single step for which a high coefficient could easily be devised. In general, however, branches have a number of steps and the 'group' coefficient is divided among them all. Genetic variation affecting any one of these is therefore likely to act on a low coefficient step with consequently much smaller differences in dominance.

It therefore appears, from our knowledge of the kinetic structure and from the experimental evidence, that the four necessary conditions for differences in dominance are rather unlikely to be met and that the rarity of observed cases is consistent with the analysis.

7.4. Appendix

Proof of the identity of the Dominance Indices of all the variables with respect to variation at one enzyme in an unsaturated system.

Using the branched system as an example, the steady state properties are fully described by the three linear rate equations $J_A = (X_A - S/K_{A,S})/a$, $J_B = (S - X_B/K_{S,B})/b$, and $J_C = (S - X_C/K_{S,C})/c$, and by (7.2). There are four consistent equations in four unknown variables J_A, J_B, J_C and S. Explicit solutions for the variables can be obtained from the matrix

$$\begin{bmatrix} aK_{A,S} & 0 & 0 & 1 \\ 0 & -b & 0 & 1 \\ 0 & 0 & -c & 1 \\ 1 & -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} J_A \\ J_B \\ J_C \\ S \end{bmatrix} = \begin{bmatrix} X_A K_{A,S} \\ X_B/K_{S,B} \\ X_C/K_{S,C} \\ 0 \end{bmatrix}$$

or $$\mathbf{Q} \quad \mathbf{z} = \mathbf{y}.$$

The solution to the column vector $\mathbf{z}$ is obtained from

$$\mathbf{z} = \mathbf{Q}^{-1}\mathbf{y}. \tag{7.15}$$

Any system of monomolecular non-saturable enzymes can be expressed in these terms. Each enzyme activity term is first order and only occurs once in the matrix Q. Consider the two allele case. E is the wild type, $E + M$ the mutant and $E + \lambda M$ the heterozygote enzyme activity (in any units). The wild type phenotypic value is $f(E)$, the mutant $f(E + M)$ and the heterozygote $f(E + \lambda M)$. In the case of a 'null' mutant, i.e. the mutant allele is a 'loss-of-function', the value of M would be $-E$, hence $E + M = 0$. The phenotypic value, $f(E + M)$ in such a case would not necessarily be equal to zero since not all characters are equally dependent on a single enzyme activity. The value of λ would, in most cases be 1/2 (*heterozygote enzyme activity* is intermediate between wild type and mutant, see e.g. Kacser and Burns, 1981; Middleton and Kacser, 1983), though of course, not the heterozygote phenotype. If we specify the three enzyme activities, the three phenotypic functions, $f(E)$, $f(E + M)$ and $f(E + \lambda M)$, can be calculated from the kinetic equations or the

matrix (7.15). The f function could, for example, be one of the branch fluxes, say, J_B. Similarly, taking another function of the same three enzyme activities (the character, say, flux J_C) we can obtain $g(E)$, $g(E + M)$ and $g(E + \lambda M)$. These functions now replace the phenotypic values in the Dominance Index definition (7.1) giving,

$$D_f = [f(E) - f(E + \lambda M)] \; / \; [f(E) - f(E + M)], \tag{7.16}$$

$$D_g = [g(E) - g(E + \lambda M)] \; / \; [g(E) - g(E + M)]. \tag{7.17}$$

Equations (7.16) and (7.17) can be expanded as a Taylor's series:

$$D_f = [f(E) - (f(E) + \lambda M f'(E) + \lambda^2 M^2 f''(E)/2! + \ldots)] \; / \; [f(E) - (f(E) + M f'(E) + M^2 f''(E)/2! + \ldots)], \tag{7.18}$$

$$D_g = [g(E) - (g(E) + \lambda M g'(E) + \lambda^2 M^2 g''(E)/2! + \ldots)] \; / \; [g(E) - (g(E) + M g'(E) + M^2 g''(E)/2! + \ldots)]. \tag{7.19}$$

From (7.15), it follows that

$$\partial \mathbf{z}/\partial E = (\partial \mathbf{Q}^{-1}/\partial E)\mathbf{y} + \mathbf{Q}^{-1}\partial \mathbf{y}/\partial E, \tag{7.20}$$

(see e.g. Graham, 1981) where E is an enzyme activity occurring in matrix $\mathbf{Q}$. Noting that the vector $\mathbf{y}$ contains no elements with enzyme activity parameters (and therefore its derivative is zero) the derivative of the inverse matrix can be re-expressed as

$$\partial z/\partial E = -\mathbf{Q}^{-1}(\partial \mathbf{Q}/\partial E)\mathbf{Q}^{-1}\mathbf{y}. \tag{7.21}$$

Similarly the second derivative is given by

$$\partial^2 z/\partial E^2 = 2\mathbf{Q}^{-1}(\partial \mathbf{Q}/\partial E)\mathbf{Q}^{-1}(\partial \mathbf{Q}/\partial E)\mathbf{Q}^{-1}\mathbf{y}. \tag{7.22}$$

Since E occurs only once and linearly in $\mathbf{Q}$, $\partial \mathbf{Q}/\partial E$ is the product of a scalar and an elementary matrix with non-zero element at the position of enzyme activity E. Turning now to equation (7.22), the term

$$(\partial \mathbf{Q}/\partial E)\mathbf{Q}^{-1}\partial \mathbf{Q}/\partial E$$

is the same elementary matrix multiplied by a different scalar. Clearly therefore, expressions (7.21) and (7.22) are proportional to one another as are higher order derivatives. Thus for any complexity of a linear system,

$$f'(E)/g'(E) = f''(E)/g''(E) = f'''(E)/g'''(E) = \ldots = \text{constant}. \tag{7.23}$$

Relation (7.23) together with equations (7.18) and (7.19) imply that,

$$D_f = D_g. \tag{7.24}$$

Therefore, for a system of any structural complexity having unsaturated enzymes, the Dominance Index for any character will be identical to that for any other character.

Chapter 8
General Discussion

In this thesis, two quite different approaches for studying quantitative variation are utilised. The basic difference between the two concerns assumptions of the mode of gene action. The first approach arbitrarily assumes that the effects of alleles are additive within and between loci (i.e. there is no dominance or epistasis at the level of the trait), and they affect fitness only through the quantitative trait. Although many characters show substantial additive genetic variation (Falconer, 1981), new mutant alleles may often have large dominance effects. This approach, however, allows a general analysis of the equilibrium and dynamic behaviour of models with linkage, mutation and selection in finite populations. The second approach assumes that the quantitative character is controlled by a particular class of genes, those controlling enzyme activities, and attempts to model a metabolic system. How might we unify these two approaches? One possible way would be to assume a model where recurrent mutation affects loci controlling enzyme activities which in turn control a metabolic system. The quantitative characters in the system would be the fluxes and substrate concentrations. It could be assumed that fitness were a function of one or several of these characters. The model of the mutation process could assume that each enzyme activity were controlled by one or several loci with the effects of new mutant alleles sampled from some distribution. This would model the range of mutational changes that can occur at a locus controlling an enzyme activity. For example, for a structural locus, some mutations would have large negative effects on enzyme activity if they occurred at the active sites; more subtle changes might be possible for changes elsewhere in the molecule. Non-additive genetic variance of flux would automatically arise. Dominance and epistasis would be greatest for alleles of largest effect. Further, models with pleiotropy could easily be investigated by assuming a branched pathway and the direction of pleiotropic gene action would depend on the enzyme's function in the system.

It is possible (but difficult in practice) to measure *in vivo* fluxes or metabolite concentrations in higher organisms and to study the effects of enzyme variation on them (e.g Flint *et al.*, 1981; Woodrow, 1986; Dykhuizen, Dean and Hartl, 1987). It is necessary, however, to contrast fluxes and pools, whose relationship to gene action is in principle fairly clear, with more complex

characters which have been commonly the subject of quantitative genetics. How closely are such complex characters related to metabolic fluxes and intermediate concentrations? They are likely to be functions of many fluxes, of systems controlling the cell cycle and cellular differentiation, and the timing of events in development and embryogenesis. Falconer (1981, Ch. 20 p.314) considers "the genes causing quantitative variation" and concludes that "it seems unlikely that they (the genes) are there solely to 'control' the character through which we recognize them. Their effects on metric characters are much more likely to be secondary to some other function... It seems most likely, therefore, that all sorts influence metric characters, no matter what the function of the gene product is."

The identification of genes causing variation in quantitative traits and elucidating their functions is a fundamental problem of quantitative genetics. For the final part of this thesis, recently developed molecular methods which enable identification of such genes will be considered. Molecular methods provide the opportunity both to induce mutations and to study their effects on traits of interest, and thus to bridge some of the gap between molecular and quantitative genetics. The technique of insertional mutagenesis which involves the induction of mutations by insertion of foreign DNA into the genome is a powerful tool for such analysis. Genetic variation induced by insertional mutagenesis is also of interest because the fixation of insertional mutations can lead to selection responses. Let us also examine some results concerning the induction of variation in quantitative traits by insertional mutagenesis and the possible use of such technology in animal improvement. In order to compare the power of insertional mutagenesis with conventional methods, as a starting point consider the induction of genetic variation by X-ray mutagenesis.

8.1. 'Conventional' artificial mutagenesis

Artificial mutagenesis has been of little use in generating new variation in commercial traits in animal species and its success in plants has been limited (Gottschalk and Wolff, 1983). Experimentally, the most extensive quantitative genetic studies are of X-ray mutagenesis of various *Drosophila melanogaster* bristle traits.

Classically, X-ray mutation rates are measured by a sex-linked lethal test in

Drosophila. Early studies showed a linear relationship between the rate of sex-linked lethal mutations and X-ray dose (reviewed by Auerbach, 1976, Ch. 5). One kR of X-irradiation induces a frequency of about 0.03 sex-linked lethals in Drosophila spermatozoa. As the average spontaneous rate to sex-linked lethals from a range of natural and laboratory stocks is about 0.0026 (reviewed by Crow and Temin, 1964), 1kR of X-rays increases the rate of sex-linked mutations about ten-fold. To compare these rates with rates of generation of variation in Drosophila bristle traits by X-rays, the experiments of Kitagawa (1967) and Hollingdale and Barker (1971) are most relevant. They provide data on response to selection of abdominal bristles from inbreds subject to X-rays each generation. Such data can be used to estimate V_M/V_E by employing theory for predicting selection responses from new mutations (Hill, 1982b). The large experiment of Clayton and Robertson (1964) cannot be used for this purpose because the X-irradiation was done during the 150 generations prior to artificial selection, so most new variation would have been lost because of drift and natural selection whose strength is unknown. Hill (1982b) derived an approximate expression for the cumulative response (C_t, in units of environmental standard deviation) to generation t from new mutations of small additive effects on the trait,

$$C_t = 2N_e i(V_M/V_E)[t - 2N_e(1\text{-}\exp(-t/2N_e))]. \qquad (8.1)$$

This can be rearranged to give V_M/V_E in terms of C_t, N_e and selection intensity (i). The population parameters, responses and estimates of V_M/V_E from Kitagawa's (1967) and Hollingdale and Barker's (1971) experiments are given in Table 8.1. Responses in the non-irradiated controls were small, but give values of V_M/V_E close to 0.001, a figure frequently quoted in the literature for this trait (see Ch. 1). If a value of V_M/V_E = 0.001 for spontaneous mutations of abdominal bristles is assumed, then 1kR of X-irradiation increases the rate at which new mutational variation is generated by a factor of about 6. It should be noted, however, that equation (8.1) assumes an 'infinitesimal' model with mutants of small selective value (Ns). The true distribution of effects of new mutants is not known, but the presence of large effects will lead (8.1) to over-estimate V_M/V_E.

Table 8.1.

Selection responses in bristle number of *D. melanogaster* in inbreds and estimates of V_M/V_E.

Reference	N_e	i	t	C_t	X-irradiation	V_M/V_E	(V_M/V_E)/kR
Kitagawa (1967)	8.4	1.4	17.5	3.36	1.5kR	0.0100	0.0067
	8.4	1.4	17.5	0.19	-	0.0006	-
Hollingdale &	140	0.8	20	1.04	1.0kR	0.0058	0.0058
Barker (1971)	140	0.8	20	0.16	-	0.0006	-

X-ray induced V_M/V_E estimated from cumulative response to selection averaged over lines using (8.1). Effective population size (N_e) was assumed to be 70% of actual size (Falconer, 1981, p.66).

In view of the relatively small responses obtained in these experiments and the number of generations required, it is perhaps not surprising that there has been little success using radiation to induce variation in vertebrates. Small improvements were, however, detected in an irradiated mouse line at an apparent body weight selection limit (Roberts, 1967).

8.2. Mutagenesis by transposable elements

The classical work on maize showed the existence of mutator loci, and similar phenomena were subsequently discovered in Drosophila. Elevated mutation rates have been shown to be caused by the movement of transposable elements, the best characterized system being the P-M system of *D. melanogaster* (see Engels, 1988, for a review). Some strains carry *P* elements, some do not, and in crosses between them a number of unusual features collectively termed "hybrid dysgenesis", are noted. In particular, mutations occur as a consequence of movement of these elements into and out of many sites in the DNA.

Recently, the movement of *P* elements following such crosses has been shown

to induce large amounts of genetic variation in bristle traits (Mackay, 1985, 1987) and in fitness (Yukuhiro *et al.*, 1985; Mackay, 1986; Fitzpatrick and Sved, 1986). Mackay's (1987) estimate of V_M/V_E per generation for abdominal bristle score was of the order of 0.1, i.e. approaching twenty times greater than that induced by 1kR of X-rays (Table 8.2). The rate of production of sex-linked lethals by *P* element transposition in "dysgenic" crosses was estimated to be 0.03 by Simmons *et al.* (1980). As discussed by Engels (1988), this rate may include mutations arising from excision as well as insertion. The rate of mutation from insertion alone was estimated to be 0.008 (Simmons et al., 1985). A more recent study (Eanes *et al.*, 1988) gives values consistent with this figure.

Table 8.2.

Comparison of spontaneous mutation rates with those from X-rays and *P* element transposition in *D. melanogaster*.

	Spontaneous	1kR X-rays	1 generation of *P* element transposition
Recessive lethal rate	0.0026 (a)	0.03 (b)	0.03 (c)
V_M/V_E for abdominal bristles	0.001 (d)	0.0062 (e)	0.11 (f)

(a) Ref. Crow and Temin (1964). (b) Ref. Auerbach (1976). (c) Presumed to include excision and rearrangements (Simmons *et al.*, 1980). (d) Average from published data (see Hill, 1982b). (e) Average from Kitagawa's (1967) and Hollingdale and Barker's (1971) results. (f) Average figure for reciprocal crosses of Mackay (1987).

Thus, Mackay's results show a greater increase in mutational variance in bristle traits than the increased mutation rate to recessive lethals observed by others. One kR of X-rays and one generation of *P* element transposition induced by a dysgenic cross both increase the rate of recessive lethals to about ten times the spontaneous rate. However, 1kR of X-rays leads to a six-fold increase in the rate of production of genetic variation of bristle number

whereas one generation of transposition leads to a hundred-fold increase. Several hypotheses are possible to explain the apparent discrepancy.

(i) The frequency distribution of mutant effects on abdominal bristles induced by X-rays and *P* element transposition is extremely leptokurtic and a few mutants of very large effect happened to be obtained by chance with *P* elements, but not X-rays. This explanation, however, seems unlikely in view of the apparent consistency of the results of independent experiments (Mackay 1985, 1986, 1987).

(ii) There were "hot spots" of *P* element activity in Mackay's lines which for some reason caused proportionally more variation in abdominal bristles and fitness traits than X-linked lethal mutations. This is probably part of the explanation for the difference in mutation rates because there is considerable variation between loci in rates of *P* element insertion (Engels, 1988). The same or similar allele was probably obtained independently in different abdominal bristle lines (Mackay, 1985). Some of these may have been caused by excision rather than insertion. Results of Eanes *et al* (1988) indicate that deletions and rearrangements of *P* elements already in place may be the most important source of mutations of large effect seen in dysgenic crosses. Intriguingly, in *Drosophila ananassae* a class of mutations, *Om*, which produce pleiotropic eye morphology defects has been shown to be caused by insertion of *tom* transposable elements at 20 different loci (Shrimpton *et al.*, 1986).

(iii) The types of mutations caused by the insertion/excision of *P* elements may be different from those caused by X-rays, and as such have more effect on quantitative traits. There is evidence that *P* elements insert preferentially in the 'control' region upstream from amino acid coding sequences (reviewed by Engels, 1988). T.F.C. Mackay (personal communication) has argued that such insertions cause changes in the 'regulation' of genes which gives the necessary range of subtle quantitative effects to generate genetic variation in quantitative traits.

8.3. Possibilities in vertebrates

The finding that large amounts of quantitative variation can be generated by insertional mutagenesis is surprising and, as discussed above, the explanation is not clear. Presumably, mutagenesis by transposition is powerful because it is

highly specific, in contrast to X-irradiation which causes non-specific damage so that doses cannot be raised sufficiently to produce much useful variation. Although transposable elements of the type found in Drosophila are not known in vertebrate species, there are other methods currently available for introducing insertional variation. An important point about such methods, which makes them different from conventional mutagenesis, is that a mutated gene is 'tagged' by the DNA insert and can subsequently cloned and its expression studied.

(i) DNA micro-injection. This method applies to mammals, and has been successful with large farm animals (Hammer *et al.*, 1985). A foreign DNA species (the particular sequence used is of little relevance for these purposes) is injected in multiple copies into the pronucleus of the fertilised egg, which can be re-introduced into a foster mother. DNA inserts apparently randomly, usually at a single site as a large tandem repeat (Palmiter and Brinster, 1986). In the mouse, the insertion of foreign DNA into the germ line to make transgenics by such methods has also been shown to induce mutations. In a survey, Palmiter and Brinster (1986) concluded that 9 visible or lethal mutations occurred in 110 transgenic mouse strains generated by DNA micro-injection.

(ii) Use of Electric Fields. DNA enters cells subjected to a strong electric field which causes temporary holes in the membrane (electroporation), with similar results to micro-injection (e.g. Chu, Hayakawa and Berg, 1987). With *iontophoresis* foreign DNA is again introduced into the pronucleus using a micro-pipette. Instead of using pressure to expel the contents as with micro-injection an electrical potential difference is set up between the pipette contents and the cell medium. DNA enters the pronucleus in a manner analogous to electrophoresis and inserts into the genome either as tandem repeats or as multiple single copies (Lo, 1983; C.W. Lo, personal communication).

(iii) Retroviral Infection. There are several routes for obtaining proviral inserts in the germ line which have been developed in the mouse (reviewed by Gridley *et al.*, 1987). The germ line can be infected with retrovirus from the 4-cell stage of the embryo to the midgestation embryo stage. Alternatively, embryonic stem (ES) cells (Evans and Kaufman, 1981) can be infected and these subsequently introduced into blastocysts to make chimaeras (Bradley *et al.*, 1984) ES cells can contribute to the germ line. Proviruses insert into the genome apparently

independently and at a large number of possible sites, the number of sites of insertion depending on the number of cycles of infection. Recent studies show a small degree of non-randomness in sites of insertion of proviruses (Shih, Stoye and Coffin, 1988). Apparently there is a subset of preferred sites, and a much larger number of less probable sites. Using ES cells, the number of inserts can be large, with tens of sites possible per cell (Robertson *et al.*, 1986). A similar frequency of mutation as for DNA micro-injection also occurred for mutagenesis by insertion of proviruses after retroviral infection of embryos, albeit with a smaller sample (Gridley *et al.*, 1987). (Larger samples of proviral integration events appear to show a mutation rate to recessive lethals closer to 1 per 40 insertions (M.J. Evans, personal communication)). The retroviruses used in these experiments are usually defective, i.e. depend on externally supplied helper functions. Such methods could, in principle, be extended to other mammals or birds.

Nothing is presently known about quantitative variation generated by insertional mutagenesis in mammals. The use of proviruses has currently the greatest potential because, like some Drosophila transposable elements, they are likely to insert near the 5' end of genes (Vijaya, Steffen and Robinson, 1986; Rohdewohld *et al.*, 1987), and multiple sites of insertion can be generated.

8.4. Potential contributions of mutations to response

The time scale required and possible benefit of increased mutation in a quantitative trait can be evaluated by simplifying (8.1) assuming $t/2N_e$ is small:

$$C_t = t^2 i(V_M/V_E)/2. \qquad (8.2)$$

The response from mutation increases approximately quadratically with generation number. Predicted responses for a range of values of V_M/V_E and two initial heritabilities assuming the 'infinitesimal' model and ignoring changes in genotypic variance from disequilibrium and selected changes in allele frequencies are shown in Table 8.3. If large quantities of useful variation could be generated, as occurred in Mackay's Drosophila lines, selection responses can be substantially increased. It should be noted, however, that this model has

several limitations, the most severe of which are: an equal frequency of beneficial and deleterious new mutations is assumed, although most mutations for traits connected with fitness are likely to be harmful; the infinitesimal model is assumed so (8.2) underestimates the possible response (Hill, 1982b).

Table 8.3.

Predicted selection responses from existing and new mutational variation in units of $i\sigma_P/\sigma_E$.

	$h^2 = 0.1$		$h^2 = 0.5$	
V_M/V_E	t=5	t=10	t=5	t=10
0	0.50	1.00	2.50	5.00
0.001	0.51	1.05	2.51	5.05
0.01	0.62	1.50	2.62	5.50
0.1	1.75	6.00	3.75	10.00

There is assumed to be negligible change in phenotypic variance due to selection on existing variation or on variation induced by mutation. An infinitesimal model of many mutants of small effect is assumed.

8.5. Isolation of genes by insertional mutagenesis

The availability of insertional mutagenesis opens up a further possible use in quantitative genetics. Rather than attempting to induce quantitative variation with a view to fixing beneficial alleles (as in the previous discussion), insertional mutagenesis might also be used to attempt to dissect the 'genetic architecture' of a quantitative character. The effects of insertion on traits of interest are likely to be large because genes may be disrupted or changed in activity or expression. Such mutations could, in principle, be identified by fixing mutant alleles in selection lines or associating specific inserts with an effect on the trait by statistical means. The particular benefit of the method is that the insert acts as a 'tag' for subsequent molecular cloning and it has been used in Drosophila to isolate *smooth*, an allele with major effect on bristle score (A.J. Leigh Brown, T.F.C. Mackay and A.E. Shrimpton, in preparation). Such

approaches are extendable, in principle, to vertebrate species. A likely problem, however, is that many alleles of large effect are recessive (Kacser and Burns, 1981) and their presence may be difficult to determine in an experimental population. Schemes involving subdivision and inbreeding are likely to improve the chance of expression and subsequent fixation of beneficial recessive alleles (Madalena and Hill, 1972).

8.6. Concluding comments

The above discussion illustrates some of the ways which molecular techniques involving insertional mutagenesis can tell us about the nature of quantitative variation. In Drosophila, where saturation of the genome with transposable elements at many sites is possible, questions can be asked concerning the distribution of effects of mutants induced by the elements on a trait, and the number of genes (or sites) which can give a detectable effect. Having isolated and cloned genes of interest, it will be possible to study variation at the molecular level contributing to variation in the trait. An important technique is the exact replacement of a gene by a cloned copy (gene transplacement) first developed in bacterial species, later in yeasts (reviewed by Smith, 1985), and there has been progress in developing this technique in the mouse (Jackson, 1988). This may be particularly powerful for analysing short regions of the genome or genes because these could be transplaced into an inbred background and the genetic variation in the trait measured.

One cannot be so optimistic about the use of molecular technology in discriminating between models of the maintenance of genetic variation. The ascertainment of effects of alleles causing variation in a trait becomes increasingly difficult as the effect of the allele becomes smaller (see Shrimpton and Robertson, 1988 for a study illustrating this point using 'conventional' methods in Drosophila). As discussed in Ch. 4 and 5, a critical question is the mode of operation of natural selection. The selectionist-neutralist controversy over the maintenance of molecular polymorphism shows that such problems do not lend themselves to simple and general solutions.

References

Abramowitz, M. and Stegun, I.A. (1965). *Handbook of Mathematical Functions.* New York: Dover publications.

Asante, E.A. (1988). Biochemical genetics of lipid metabolism in chickens and mice. Ph.D. Thesis, Edinburgh University.

Asante, E.A. and Bulfield, G. (1988). Activities of NADPH-generating enzymes in genetically fat and lean chickens. In *Leanness in Domestic Birds,* ed. B. Leclercq and C.C. Whitehead. London: Butterworths.

Atkins, K.D. and Thompson, R. (1986). Predicted and realised responses to selection for an index of bone length and body weight in Scottish Blackface sheep. *Animal Production 43,* 421-435.

Auerbach, C. (1976). *Mutation research.* London: Chapman and Hall.

Avery, P.J. and Hill, W.G. (1977). Variability in genetic parameters among small populations. *Genetical Research 29,* 193-213.

Avery, P.J. and Hill, W.G. (1979). Variance in quantitative traits due to linked dominant genes and variance in heterozygosity in small populations. *Genetics 91,* 817-844.

Bailey, D.W. (1959). Rates of subline divergence in highly inbred strains of mice. *Journal of Heredity 50,* 26-30.

Bannister, D.W., Lee., A., Whitehead, C. and Griffin, H. (1984). Lipogenic enzyme activity and fructose 2,6-bisphosphate concentrations in livers of two lines of domestic fowl (*Gallus domesticus*) selected for different body fat content. *International Journal of Biochemistry 16,* 1301-1305.

Barton, N.H. (1986). The maintenance of polygenic variation through a balance between mutation and stabilizing selection. *Genetical Research 47,* 209-216.

Barton, N.H. and Turelli, M. (1987). Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genetical Research 49,* 157-173.

Bonhomme, F., Guenet, J-L., Dod, B., Moriwaki, K. and Bulfield, G. (1987). The polyphyletic origin of laboratory inbred mice and their rate of evolution.

Biological Journal of the Linnean Society 30, 51–58.

Box, G.E.P. and Muller, M.E. (1958). A note on generation of normal deviates. *Annals of Mathematical Statistics 28,* 610–611.

Bradley, A., Evans, E., Kaufman, M.H. and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature 309,* 255–256.

Bulfield, G. and Kacser, H. (1974). Histidinaemia in mouse and man. *Archives of Disease in Childhood 49,* 545–552.

Bulfield, G., Moore, E.A., and Kacser, H. (1978). Genetic variation in activity of the enzymes of glycolysis and gluconeogenesis between inbred strains of mice. *Genetics 89,* 551–561.

Bulmer, M.G. (1971). The effect of selection on genetic variability. *American Naturalist 105,* 201–211.

Bulmer, M.G. (1972). The genetic variability of polygenic characters under optimizing selection, mutation and drift. *Genetical Research 19,* 17–25.

Bulmer, M.G. (1976). The effect of selection on genetic variability: a simulation study. *Genetical Research 28,* 101–117.

Bulmer, M.G. (1980). *The Mathematical Theory of Quantitative Genetics.* Oxford: Clarendon Press.

Bulmer, M.G. (1988). Maintenance of genetic variability by mutation-selection balance: a child's guide through the jungle. In *Proceedings of The XVIth International Congress of Genetics,* in press.

Burger, R. (1986). On the maintenance of genetic variation: Global analysis of Kimura's continuum of alleles model. *Journal of Mathematical Biology 24,* 341–351.

Burger, R.E., Wagner, G.P. and Stettinger, F. (1988). How much heritable variation can be maintained in finite populations by a mutation selection balance? *Evolution* in press.

Burns, J.A. and Kacser, H. (1977). Allosteric Repression: An Analysis. *Journal of*

Theoretical Biology 68, 199–213.

Chakraboty, R. and Nei, M. (1982). Genetic differentiation of quantitative characters between populations or species. I. Mutation and random genetic drift. *Genetical Research 39,* 303–314.

Chu, G., Hayakawa, H. and Berg, P. (1987). Electroporation for the efficient transfection of mammalian cells with DNA. *Nucleic Acids Research 15,* 1311–1326.

Clarke, B. (1972). Density-dependent selection. *American Naturalist 106,* 1–13.

Clarke, B. and O'Donald, P. (1964). Frequency-dependent selection. *Heredity 19,* 201–206.

Clayton, G. and Robertson, A. (1955). Mutation and quantitative variation. *American Naturalist 89,* 151–158.

Clayton, G.A., Morris, J.A. and Robertson, A. (1957). An experimental check on quantitative genetical theory. I. Short-term responses to selection. *Journal of Genetics 55,* 131–151.

Clayton, G.A. and Robertson, A. (1964) The effects of X-rays on quantitative characters. *Genetical Research 5,* 410–422.

Cleland, W.W. (1963). The kinetics of enzyme catalysed reactions with two or more substrates or products. I. Nomenclature and rate equations. *Biochimica et Biophysica Acta 67,* 104–137.

Cockerham, C.C. and Tachida, H. (1987). Evolution and maintenance of quantitative genetic variation by mutations. *Proceedings of the National Academy of Sciences, USA 84,* 6205–6209.

Cornish-Bowden, A. (1987). Dominance is not inevitable. *Journal of Theoretical Biology 125,* 333–337.

Crow, J.F. and Kimura, M. (1964). The theory of genetic loads. *Proceedings of the XIth International Congress of Genetics 2,* 495–505.

Crow, J.F. and Kimura, M. (1970). *An Introduction to Population Genetics Theory.* New York: Harper and Row.

Crow, J.F. and Temin, R.G. (1964). Evidence for the partial dominance of recessive lethal genes in natural populations of Drosophila. *American Naturalist 98*, 21–33.

David, J.R. and Bocquet, C. (1975). Similarities and differences in latitudinal adaptation of two Drosophila sibling species. *Nature 257*, 588–590.

Dean, A.M., Dykhuizen, D.E. and Hartl, D.L. (1986). Fitness as a function of β-galactosidase activity in *Escherichia coli*. *Genetical Research 48*, 1–8.

Dudley, J.W. (1977). 76 generations of selection for oil and protein percentage in maize. In *Proceedings of the International Conference on Quantitative Genetics* (ed. E. Pollak, O. Kempthorne and T.B. Bailey, Jr.), pp.459–473. Ames: Iowa State University Press.

Dykhuizen, D.E., Dean, A.M. and Hartl, D.L. (1987). Metabolic Flux and Fitness. *Genetics 115*, 25–31.

Eanes, W.F., Wesley, C., Hey, J., Houle, D., Ajioka, J.W. (1988). The fitness consequences of *P* element insertion in *Drosophila melanogaster*. *Genetical Research 52*, 17–26.

Enfield, F.D. (1980). Long term effects of selection; the limits to response. In *Selection Experiments in Laboratory and Domestic Animals* (ed. A. Robertson), pp.69–86. Slough: Commonwealth Agricultural Bureaux.

Engels, W.R. (1988). *P* elements in Drosophila. In *Mobile DNA* (ed. D.E. Berg and M.M. Howe) (in press).

Evans, M.J. and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature 292*, 154–156.

Ewens, W.J. (1979). *Mathematical Population Genetics*. Berlin: Springer-Verlag.

Falconer, D.S. (1973). Replicated selection for body weight in mice. *Genetical Research 22*, 291–321.

Falconer, D.S. (1981). *Introduction to Quantitative Genetics*, 2nd edn. London: Longman.

Fell, A. and Sauro, H.M. (1985). Metabolic control and its analysis. Additional

relationships between elasticities and control coefficients. *European Journal of Biochemistry 148,* 555-561.

Felsenstein, J. (1965). The effect of linkage on directional selection. *Genetics 52,* 349-363.

Festing, M. (1973). A multivariate analysis of subline divergence in the shape of the mandible in C57BL/Gr mice. *Genetical Research 21,* 121-132.

Fisher, R.A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh, 52,* 399-433.

Fisher, R.A. (1928). The Possible Modification of the Response of the Wild Type to Recurrent Mutations. *American Naturalist 62,* 115-126.

Fitzpatrick, B.J. and Sved, J.A. (1986). High levels of fitness modifiers induced by hybrid dysgenesis in *Drosophila melanogaster. Genetical Research 48,* 89-94.

Flint, H.J., Porteous, D.J. and Kacser, H. (1980). Control of the Flux in the Arginine Pathway of *Neurospora crassa.* The flux from Citrulline to Arginine. *Biochemical Journal 190,* 1-15.

Flint, H.J., Tateson, R.W., Barthelmess, I.B., Porteous, D.J., Donachie, W.D. and Kacser, H. (1981). Control of the Flux in the Arginine Pathway of *Neurospora crassa.* Modulation of enzyme activity and concentration. *Biochemical Journal 200,* 231-246.

Frankham, R., Jones, L.P. and Barker, J.S.F. (1968). The effects of population size and selection intensity in selection for a quantitative character in Drosophila. I. Short term response to selection. *Genetical Research 12,* 237-248.

Gibson, J.B. and Bradley, B.P. (1974). Stabilising selection in constant and fluctuating environments. *Heredity 33,* 293-302.

Gillespie, J.H. (1984). Pleiotropic overdominance and the maintenance of genetic variation in polygenic characters. *Genetics 107,* 321-330.

Gottschalk, W. and Wolff, G. (1983). *Induced Mutations in Plant Breeding.* Berlin: Springer-Verlag.

Graf, J-D. and Ayala, F.J. (1986). Genetic variation for superoxide dismutase level in *Drosophila melanogaster. Biochemical Genetics 24,* 153-168.

Graham, A. (1981). *Kronecker Products and Matrix Calculus with Applications.* Chichester: Ellis Horwood.

Grewal, M.S. (1962). The rate of genetic divergence of sublines in the C57BL strain of mice. *Genetical Research 3,* 226-237.

Gridley, T., Soriano, P. and Jaenisch, R. (1987). Insertional mutagenesis in mice. *Trends in Genetics 3,* 162-166.

Groen, A.K., Van Der Merwe, R., Westerhoff, H.V. Wanders, H.V., Wanders, R.J.A., Ackerbroom, T.P.M. and Tager, J.M. (1982). Control of Metabolic Fluxes pp.9-37. In: Metabolic Compartmentation, *Ed.* H. Fries. New York: Academic Press.

Guenet, J.-L., Marchal, G., Milon, G., Tambourin, P. and Wendley, F. (1979). Fertile dominant spotting in the house mouse. *Journal of Heredity 70* 9-12.

Haldane, J.B.S. (1919). The combination of linkage values and the calculation of distance between loci of linked factors. *Journal of Genetics 8,* 299-309.

Haldane, J.B.S. (1932). *The Causes of Evolution.* London: Longmans.

Haldane, J.B.S. (1954). The measurement of natural selection. In Proceedings of the 9th International Congress of Genetics. (*Carylogia, Supplement to Volume 6,* Vol. 1 pp 480-487.

Hammer, R.E., Pursel, V.G., Rexroad Jr., C.E., Wall, R.J., Bolt, D.J., Ebert, K.M., Palmiter, R.D. and R.L. Brinster (1985). Production of transgenic rabbits, sheep and pigs by microinjection. *Nature 315,* 680-683.

Harris, H. (1966). Enzyme Polymorphism in Man. *Proceedings of the Royal Society of London, Series B 164,* 298-310.

Harris, H. (1980). *The Principles of Human Biochemical Genetics.* Third Edition. Amsterdam: Elsevier.

Hartl, D.L., Dykhuizen, D.E. and Dean, A.M. (1985). Limits of adaptation: the evolution of selective neutrality. *Genetics 111,* 655-674.

Heinrich, R. and Rapoport, T.A. (1974). A linear steady state treatment of enzyme chains. *European Journal of Biochemistry 42*, 89–102.

Hess, B. (1973). Organization of glycolysis: oscillatory and stationary control. *Symposium of the Society for Experimental Biology 27*, 105–132.

Hill, W.G. (1982a). Rates of change in quantitative traits from fixation of new mutations. *Proceedings of the National Academy of Sciences, USA 79*, 142–145.

Hill, W.G. (1982b). Predictions of response to artificial selection from new mutations. *Genetical Research 40*, 255–278.

Hill, W.G. and Keightley, P.D. (1988). Interrelations of mutation, population size, artificial and natural selection. In *Proceedings of the Second International Conference on Quantitative Genetics* (ed. B.S. Weir, E.J. Eisen, M.M. Goodman and G. Namkoong). Sunderland, Massachusetts: Sinauer.

Hill, W.G. and Rasbash, J. (1986). Models of long term artificial selection in finite populations. *Genetical Research 48*, 41–50.

Hill, W.G. and Robertson A. (1966). The effects of linkage on limits to artificial selection. *Genetical Research 8*, 269–294.

Hofmeyer, J.H.S., Kacser, H. and Van Der Merwe, K.J. (1986). Metabolic control analysis of moiety-conserved cycles. *European Journal of Biochemistry 155*, 631–641.

Hollingdale, B. (1971). Analysis of some genes from abdominal bristle number selection lines in *Drosophila melanogaster*. *Theoretical and Applied Genetics 41*, 292–301.

Hollingdale, B. and Barker, J.S.F. (1971). Selection for increased abdominal bristle number in *Drosophila melanogaster* with concurrent irradiation. I. Populations derived from an inbred line. *Theoretical and Applied Genetics 41*, 208–215.

Hyytia, P., Capy, P., David, J.R. and Singh, R.S. (1985). Enzymatic and quantitative variation in European and African populations of *Drosophila simulans*. *Heredity 54*, 209–217.

Jackson, I.J. (1988). No Mickey Mouse genetics. *Trends in Genetics 4*, 273–274.

Jenkins, N.A. and Copeland, N.G. (1985). High frequency germline acquisition of ecotropic MuLV proviruses in SWR/J-RF/J hybrid mice. *Cell 43*, 811-819.

Johnson, W.G. and Hong, J.L. (1986). Variation in alpha-L-fucosidase properties among 28 inbred mouse strains: six strains have high enzyme activity and heat-stabile enzyme with a variant pH-activity curve; twenty-two strains have low activity and heat-labile enzyme. *Biochemical Genetics 24*, 469-483.

Johnson, W.G., Hong, J.L. and Knights, S.M. (1986). Variation in ten lysosomal hydrolase enzyme activities in inbred mouse strains. *Biochemical Genetics 24*, 891-909.

Kacser, H. (1983). The control of enzyme systems *in vivo*: elasticity analysis of the steady state. *Biochemical Society Transactions 11*, 35-40.

Kacser, H. (1987). Dominance not inevitable but very likely. *Journal of Theoretical Biology 126*, 505-506.

Kacser, H. and Burns, J.A. (1973). The control of flux. *Symposium of the Society for Experimental Biology 27*, 65-104.

Kacser, H. and Burns, J.A. (1981). The molecular basis of dominance. *Genetics 97*, 639-666.

Kacser, H. and Porteous, J.W. (1987). Control of metabolism: what do we have to measure? *Trends in Biochemical Sciences 12*, 5-14.

Kaufman, P.K., Enfield, F.D. and Comstock, R.E. (1977). Stabilizing selection for pupal weight in *Tribolium castaneum*. *Genetics 87*, 327-341.

Keightley, P.D. and Hill, W.G. (1983). Effects of linkage on response to directional selection from new mutations. *Genetical Research 42*, 193-206.

Kemeny, J.F. and Snell, J.L. (1960). *Finite Markov Chains*. Princeton: Van Nostrand.

Kennedy, W.J. and Gentle, J.E. (1980). *Statistical Computing*. New York: Marcel Dekker Inc.

Kimura, M. (1964). Diffusion models in population genetics. *Journal of Applied Probability 1*, 177-232.

Kimura, M. (1965). A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proceedings of the National Academy of Sciences, USA 54*, 731-736.

Kimura, M. (1969). The number of heterozygous nucleotide sites maintained in a finite population due to a steady flux of mutations. *Genetics 61*, 893-903.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press.

Kingman, J.F.C. (1978). A simple model for the balance between selection and mutation. *Journal of Applied Probability 15*, 1-12.

Kinghorn, B.P. (1987). The nature of 2-locus epistatic interactions in animals: evidence from Sewall Wright's guinea pig data. *Theoretical and Applied Genetics 73*, 595-604.

Kitagawa, O. (1967). The effects of X-ray irradiation on selection response in *Drosophila melanogaster*. *Japanese Journal of Genetics 42*, 121-137.

Knox, W.E. and Messinger, E. (1958). The detection in the heterozygote of the metabolic effect of the recessive gene for Phenylketonuria. *American Journal Human Genetics 10*, 53-60.

Knuth, D.E. (1968). *The Art of Computer Programming, Vol 1: Fundamental Algorithms*. Reading, Massachusetts: Addison-Wesley.

Kojima, K. (1959). Role of epistasis and overdominance in stability of equilibria with selection. *Proceedings of the National Academy of Sciences USA 45*, 984-989.

Lack, D. (1968). *Ecological Adaptation for Breeding in Birds*. London: Methuen.

Lande, R. (1976). The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genetical Research 26*, 221-235.

Lande, R. (1980). The genetic covariance of characters maintained by pleiotropic mutations. *Genetics 94*, 203-215.

Laporte, D.C, Walsh, K. and Koshland, D.E. (1984). The branch point effect. *Journal of Biological Chemistry 259*, 14068-14075.

Latter, B.D.H. (1960). Natural selection for an intermediate optimum. *Australian Journal of Biological Science 13,* 30-35.

Latter, B.D.H. (1970). Selection in finite populations with multiple alleles. II. Centripetal selection, mutation and isoallelic variation. Genetics *66,* 165-186.

Laurie-Ahlberg, C.C. (1982). Genetic, ontogenetic and tissue-specific variation of dipeptidases in *Drosophila melanogaster. Biochemical Genetics 20,* 407-424.

Laurie-Ahlberg, C.C., Maroni, G., Bewley, G.C, Lucchesi, J.C. and Weir B.S. (1980). Quantitative genetic variation of enzyme activities in natural populations of *Drosophila melanogaster. Proceedings of the National Academy of Sciences, USA 77,* 1073-1077.

Lewontin, R.C. (1974). *The Genetic Basis of Evolutionary Change.* New York: Columbia University Press.

Lewontin, R.C. and Hubby, J.L. (1966). A molecular approach to the study of the genic heterozygosity in natural populations, II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura. Genetics 54,* 595-609.

Lindsley, D.L. and Grell, E.H. (1968). Genetic variations of *Drosophila Melanogaster. Carnegie Institute of Washington Publications 627.*

Lo, C.W. (1983). Transformation by iontophoretic microinjection of DNA: multiple integration without tandem insertions. *Molecular and Cellular Biology 3,* 1803-1814.

Lynch, M. (1988). The rate of polygenic mutation. *Genetical Research 51,* 137-148.

Lynch, M. and Hill, W.G. (1986). Phenotypic evolution by neutral mutation. *Evolution 40,* 915-935.

Mackay, T.F.C. (1981). Genetic variation in varying environments. *Genetical Research 37,* 79-93.

Mackay, T.F.C. (1985). Transposable element-induced response to artificial selection in *Drosophila melanogaster. Genetics 111,* 351-374.

Mackay, T.F.C. (1986). Transposable element-induced fitness mutations in *Drosophila melanogaster. Genetical Research 48,* 77-87.

Mackay, T.F.C. (1987). Transposable element-induced polygenic mutations in *Drosophila melanogaster. Genetical Research 49,* 225-233.

McClintock, B. (1950). The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences USA 36,* 344-355.

Madalena, F.E. and Hill, W.G. (1972). Population structure in artificial selection programmes: simulation studies. *Genetical Research 20,* 75-99.

Mather, K. and Wigan, L.G. (1942). The selection of invisible mutations. *Proceedings of the Royal Society of London B 131,* 50-64.

Middleton, R.J. and Kacser, H. (1983). Enzyme variation, metabolic flux and fitness: alcohol dehydrogenase in *Drosophila melanogaster. Genetics 105,* 633-650.

Miyashita, N. and Laurie-Ahlberg, C.C. (1986). Developmental variation in effects of the second and third chromosomes on the activities of the glucose 6-phosphate and 6-phosphogluconate dehydrogenases in *Drosophila melanogaster. Biochemical Genetics 24,* 447-467.

Mousseau, T.A. and Roff, D.A. (1987). Natural selection and the heritability of fitness components. *Heredity 59,* 181-197.

Mukai, T., Chigusa, S.I., Mettler, L.E. and Crow, J.F. (1972). Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster. Genetics 72,* 333-355.

Mukai, T. and Cockerham, C.C. (1977). Spontaneous mutation rates at enzyme loci in *Drosophila melanogaster. Proceedings of the National Academy of Sciences, USA 74,* 2514-2517.

Muller, H.J. (1964). The relation of recombination to mutational advance. *Mutation Research 1,* 2-9.

Nei, M. (1987). *Molecular Evolutionary Genetics.* New York: Columbia Universtiy Press.

Paigen, K. (1979). Acid hydrolases as models of genetic control. *Annual Review of Genetics 13*, 417–466.

Palmiter, R.D. and Brinster, R.L. (1986). Germ-line transformation of mice. *Annual Review of Genetics 20*, 465–499.

Provine, W.B. (1971). *The Origins of Theoretical Population Genetics.* The University of Chicago Press.

Rohdewohld, H., Weiher, H., Reik, W., Jaenisch, R. and Breinde, M. (1987). Retrovirus integration and chromatin structure: Moloney murine leukemia proviral integration sites map near DNase 1-hypersensitive sites. *Journal of Virology 61*, 336–343.

Roberts, R.C. (1967). The limits to artificial selection for body weight in the mouse. IV. Sources of new genetic variance - irradiation and outcrossing. *Genetical Research 9*, 87–98.

Robertson, A. (1956). The effect of selection against extreme deviants based on deviation or on homozygosis. *Journal of Genetics 54*, 236–248.

Robertson, A. (1959). The sampling variance of the genetic correlation coefficient. *Biometrics 15*, 469–485.

Robertson, A. (1962). Selection for heterozygotes in small populations. *Genetics 47*, 1291–1300.

Robertson, A. (1967). The nature of quantitative genetic variation. pp. 265–280. In: *Heritage from Mendel* (Ed. R.A. Brink). Madison, Milwaukee and London: University of Wisconsin Press.

Robertson, A. (1970). A theory of limits in artificial selection with many linked loci. In *Mathematical Topics in Population Genetics* (ed. K.Kojima) pp. 246–288. Berlin: Springer–Verlag.

Robertson, A. (1973). The validity of the optimum model. *Advances in Applied Probability 6*, 17–18.

Robertson, A. (1977). Artificial selection with a large number of linked loci. In *Proceedings of the International Conference on Quantitative Genetics* (ed. E.Pollak, O.Kempthorne and T.B.Bailey Jr.) pp 307–322. Ames: Iowa State University Press.

Robertson, A. (1978). The time of detection of recessive visible genes in small populations. *Genetical Research 31,* 255-264.

Robertson, E., Bradley, A., Kuehn, M. and M. Evans, M. (1986). Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. *Nature 323,* 445-448.

Roff, D.A. and Mousseau, T.A. (1987). Quantitative genetics and fitness: lessons from Drosophila. *Heredity 58,* 103-118.

Rose, M.R. (1982). Antagonistic pleiotropy, dominance and genetic variation. *Heredity 48,* 63-78.

Ruano, R.G, Orozco, F. and Lopez-Fanjul, C. (1975). The effect of different selection intensities on selection response in egg-laying of *Tribolium castaneum. Genetical Research 25,* 17-27.

Russell, W.A., Sprague, G.F. and Penny, L.H. (1963). Mutations affecting quantitative characters in long-time inbred lines of maize. *Crop Science 3,* 175-178.

Salter, M., Knowles, R.G. and Pogson, C.I. (1986). Quantitation of the importance of individual steps in the control of aromatic amino acid metabolism. *Biochemical Journal 234,* 635-647.

Savageau, M.A. (1969). Biochemical systems analysis I. Some mathematical properties of the rate law for the component enzymatic reactions. *Journal of Theoretical Biology 25,* 365-369.

Savageau, M.A., Voit, E.O. and Irvine, D.H. (1987). Biochemical systems theory and metabolic control theory: 1. Fundamental similarities and differences. *Mathematical Biosciences 86,* 127-148.

Shih, C.-C., Stoye, J.P. and Coffin, J.M. (1988). Highly preferred targets for retrovirus integration. *Cell 53,* 531-537.

Shrimpton, A.E., Montgomery, E.A. and Langley, C.H. (1986). *Om* mutations in *Drosophila ananassae* are linked to insertions of a transposable element. *Genetics 114,* 125-135.

Shrimpton, A.E. and Robertson, A. (1988). The isolation of polygenic factors

controlling bristle score in *Drosophila melanogaster*. II. Distribution of third chromosome bristle effects within chromosome sections. *Genetics 118*, 445-459.

Simmons, M.J., Johnson, N.A., Fahey, T.M., Nellett, S.M. and Raymond, J.D. (1980). High mutability in male hybrids of *Drosophila melanogaster*. *Genetics 96*, 479-490.

Simmons, M.J., Raymond, J.D., Laverty, T.R., Doll, R.F., Raymond, N.C., Kocur, G.C. and Drier, G.A. (1985). Chromosomal effects on mutability in the P-M system of hybrid dysgenesis in *Drosophila melanogaster*. *Genetics 111*, 869-884.

Slatkin, M. (1987). Heritable variation and heterozygosity under a balance between mutations and stabilizing selection. *Genetical Research, 50*, 53-62.

Smith, C. (1984). Rates of genetic change in farm livestock. *Research and Development in Agriculture 1*, 79-85.

Smith, M. (1985). In vitro mutagenesis. *Annual Review of Genetics 19*, 423-462.

Sprague, G.F., Russell, W.A. and Penny, L.H. (1960). Mutations affecting quantitative traits in selfed progeny of doubled monoploid maize stocks. *Genetics 45*, 855-866.

Stanbury, J.B., Wyngaarden, J.B., Fredrickson, D.S., Goldstein, J.L. and Brown, M.S. (1983). *The Metabolic Basis of Inherited Disease*. McGraw-Hill: New York.

Standal, N. and Vangen, O. (1980). Physiological effects of selection for growth rate and backfat thickness. In *Selection Experiments in Laboratory and Domestic Animals* (Ed. A. Robertson), pp. 125-130. Slough: Commonwealth Agricultural Bureaux.

Stephens, F. and Murison, J. (1982). *The IMP80 Language*. Edinburgh Regional Computing Centre.

Stuart, F., Porteous, D.J., Flint, H.,J. and Kacser, H. (1986). Control of flux in the Arginine pathway of *Neurospora crassa*: Effects of co-ordinate changes of enzyme concentration. *Journal of General Microbiology 132*, 1159-1166.

Thoday, J.M., Gibson, J.B. and Spickett, S.G. (1964). Regular responses to selection. 2. Recombination and accelerated response. *Genetical Research 5*,

1-19.

Turelli, M. (1984). Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. *Theoretical Population Biology 25,* 138-193.

Turelli, M. (1985). Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits. *Genetics 111,* 165-195.

Vijaya, S., Steffen, D.L. and Robinson, H.L. (1986). Acceptor sites for retroviral integrations map near DNase 1-hypersensitive sites in chromatin. *Journal of Virology 60,* 683-692.

Weatherall, D.J. and Clegg, J.B. (1981). *The Thallasaemia Syndromes.* Oxford: Blackwell Scientific Publications.

Webb, A.J. (1981). The Halothane Sensitivity Test. In *Proceedings of Symposium on Porcine Stress and Meat Quality.* (eds. T. Froystein, E. Slinde, N. Standal) pp.105-124. Norwegian Agricultural Food Research Society, As.

Weinberg, W. (1909a). Ueber Vererbungsgesetze beim Menschen. 1. Allgemeiner Teil. *Zeitshrift fur Induktive Abstammungs und Vererbungslehre 1,* 377-392.

Weinberg, W. (1909b). Ueber Vererbungsgesetze beim Menschen. 2. Spezieller Teil. *Zeitshrift fur Induktive Abstammungs und Vererbungslehre 2,* 276-300.

Weinberg, W. (1910). Weitere Beitrage zur Theorie der Vererbung. *Archiv fur Rassen und Gesellschafts Biologie 7,* 35-49, 169-173.

Whitehead, C.C., Hood, R.L., Heard, G.S. and Pym, R.A.E. (1984). Comparison of plasma very low density lipoproteins and lipogenic enzymes as predictors of fat content and food conversion efficiency in selected lines of broiler chickens. *British Poultry Science 25,* 277-286.

Woodrow, I.E. (1986). Control of the rate of photosynthetic carbon dioxide fixation. *Biochimica et Biophysica Acta 851,* 181-192.

Wright, S. (1921). Systems of mating. *Genetics 6,* 111-178.

Wright, S. (1922). The effects of inbreeding and crossbreeding of guinea pigs. 3. Crosses between highly inbred families. *USDA Bulletin, No. 1121.*

Wright, S. (1934). Molecular and evolutionary theories of dominance. *American Naturalist 68,* 24-53.

Yoo, B.H. (1980). Long-term selection for a quantitative character in large replicate populations of *Drosophila melanogaster. Genetical Research 35,* 1-17.

Yukuhiro, K., Harada, K. and Mukai, T. (1985). Viability mutations induced by the *P* elements in *Drosophila melanogaster. Japanese Journal of Genetics 60,* 531-537.

Zeng, Z-B. (1987). Genotypic distribution at the limits to natural and artificial selection with mutation. *Theoretical Population Biology 32,* 90-113.

Appendix: List of Publications

Keightley, P.D. and Kacser, H. (1987). Dominance, pleiotropy and metabolic structure. *Genetics 117*, 319–329.

Keightley, P.D. and Hill, W.G. (1987). Directional selection and variation in finite populations. *Genetics 117*, 573–582.

Hill, W.G. and Keightley, P.D. (1988). Interrelations of mutation, population size, artificial and natural selection. In *Proceedings of the Second International Conference on Quantitative Genetics* (ed. B.S. Weir, E.J. Eisen, M.M. Goodman, and G. Namkoong), pp 57–70. Sunderland, Massachussets: Sinauer.

Keightley, P.D. and Hill, W.G. (1988). Quantitative genetic variability maintained by mutation–stabilizing selection balance in finite populations. *Genetical Research 52*, 33–43.

Hill, W.G. and Keightley, P.D. (1988). Interaction between molecular and quantitative genetics. In *Advances in Animal Breeding.* Proceedings of symposium in honour of Professor R.D. Politiek. (ed. S. Korver, H.A.M. van der Steen, J.A.M. van Arendonk, H. Bakker, E.W. Brascamp and J. Dommerholt.) pp 41–55. Wageningen: Pudoc.

Keightley, P.D. (1988). Models of Quantitative Variation of Flux in Metabolic Pathways. *Genetics*, in press.

Keightley, P.D. and Hill, W.G. (1988). Directional Selection on Quantitative Traits Previously Under Stabilizing Selection. *Genetical Research*, accepted.

Dominance, Pleiotropy and Metabolic Structure

Peter D. Keightley and Henrik Kacser

Department of Genetics, University of Edinburgh, Edinburgh EH9 3JN, Scotland

Manuscript received February 10, 1987
Revised copy accepted June 19, 1987

ABSTRACT

It is a common observation that most mutants have similar dominance relations for all the characters they are known to affect. As a model of pleiotropic effects we investigate a branched pathway where the two outputs represent two characters whose variation is affected by changes in any of the genetically specified enzymes in the system. We consider the effects on the phenotype (fluxes or intermediate metabolites) of substitutions at one locus represented by enzyme activities of the two homozygotes (mutant and wild type) and that of the heterozygote. Dominance indices for the characters pleiotropically connected by the metabolic system are calculated. We show that if enzymes behave 'linearly,' (first order), that is if saturation and feedback inhibition or other nonlinearities are absent, all fluxes and pools have identical dominance relations. The presence of such nonlinearity, however, leads to differences in dominance between different characters and we define the conditions where such differences can be important.

THE characters which geneticists measure must arise from the physiological and metabolic processes which occur within the organism. Genetic variation results in variation of the gene products of which many are catalytic proteins, *i.e.*, enzymes. Acting through the metabolism, enzymes 'control' the variables, *i.e.*, fluxes and metabolic pools, in a quantitative manner. Metabolism is often represented by the metabolic map. This map shows us how the enzymes are connected to one another by the metabolites they share. The substrate for one reaction is, in general, the product of another. Other interactions are from metabolites that act as effectors, positive or negative, for particular enzymes. The map does not, however, define the kinetic structure of the metabolic system, that is, give us any information on the rates at which substrates are converted from one to another. The measurement of these rates—the metabolic fluxes—and of the metabolite pool concentrations is the proper subject of 'quantitative metabolism.' The methodology of quantitative metabolism is to study the effects of varying the *parameters* of the system (*e.g.*, enzyme activities) on *variables* which may be measured (*i.e.*, fluxes and metabolite pool levels). Theoretical treatments of quantitative metabolism (KACSER and BURNS 1973; HEINRICH and RAPOPORT 1974; KACSER 1983; FELL and SAURO 1985; HOFMEYER, KACSER and VAN DER MERWE 1986) have provided expectations of the behavior of living systems (*e.g.*, FLINT, PORTEOUS and KACSER 1980; FLINT *et al.* 1981; GROEN *et al.*, 1982; MIDDLETON and KACSER 1983; STUART *et al.* 1986; WOODROW 1986; SALTER, KNOWLES and POGSON 1986; DYKHUIZEN, DEAN and HARTL 1987). Because, in principle, all enzymes affect all variables and since enzyme activities are under genetic control, fluxes and pool levels are quantitative characters or very closely related to them. In living systems the values of these characters are determined by the alleles which control enzyme activities and by the environment which controls the inputs and the external effectors to the metabolic system.

In a previous paper (KACSER and BURNS 1981), the methods of quantitative metabolism were applied to the question of the effects of finite changes in enzyme activity generated by allelic differences and led to a general analysis of dominance relationships. It was shown that there is a nonlinear relationship between flux or metabolite concentration and enzyme activity. The general expectation that "null" mutants at enzyme loci are 'recessive' is explainable in these terms without necessitating an evolutionary hypothesis of "modifiers" first proposed by FISHER (1928) (see *e.g.*, MIDDLETON and KACSER 1983; DEAN, DYKHUIZEN and HARTL 1986; CORNISH-BOWDEN 1987; KACSER 1987). On the other hand, small differences in the enzyme parameters of mutant and wild type were shown to result in an intermediate heterozygote phenotype. Here, we extend the KACSER and BURNS (1981) treatment to the problem of pleiotropic effects of enzyme variation.

Since the metabolic system is highly interactive, genetic variation at one locus will in principle affect all the characters. This, together with the interactive nature of development gives us a general expectation of pleiotropy. It does not, however, imply that any genetic variation affects all the characters in the same way. Intuitively, we would expect some characters to be 'close' to one another and others to be more

an arbitrary number of first order enzymes in each.

By solving the simultaneous equations, expressions for each of the fluxes J_A, J_B and J_C can be obtained in which the enzyme activities appear as parameters. Although a character, such as a flux or a pool, is a function of all the enzyme activities in the system, we are interested in the effect of variation in one enzyme only. This can be anywhere in the system. We consider the two-allele case: wild type, mutant and heterozygote for which three different values of one of the enzyme activities will apply. By inserting these three different enzyme values in each of the equations for J_A, J_B and J_C, the phenotypic values for the fluxes are obtained. Finally, insertion of these in to the expression for the dominance index (1) allows us to compare how changes at one locus affects the dominance relations for the pleiotropically related fluxes.

In Appendix 2 we give a general proof that for a system of unsaturated enzymes, of any complexity, the dominance index with respect to any one locus is exactly the same for all characters (fluxes or metabolite pools) affected by the allelic substitutions. Although the dominance indices are identical, the measured differences in the characters may be very different. If, *e.g.*, we find 'recessivity' in one character, say a flux, ($D \rightarrow 0$), then we shall find the same in another pleiotropically related character, say a pool. The underlying feature which generates these identities is that in a system with linear equations for each step, all the fluxes are linearly related to all the pools. This is illustrated in Figure 2 which is shown as a 'reflection diagram.' (See BURNS and KACSER 1977.) Such a diagram shows the effects of the independent variable (enzyme activity A in this case) on a dependent variable (metabolite S in this case) which, in turn, affects further dependent variables (the three fluxes). The decomposition into the functional components aids the understanding of the system.

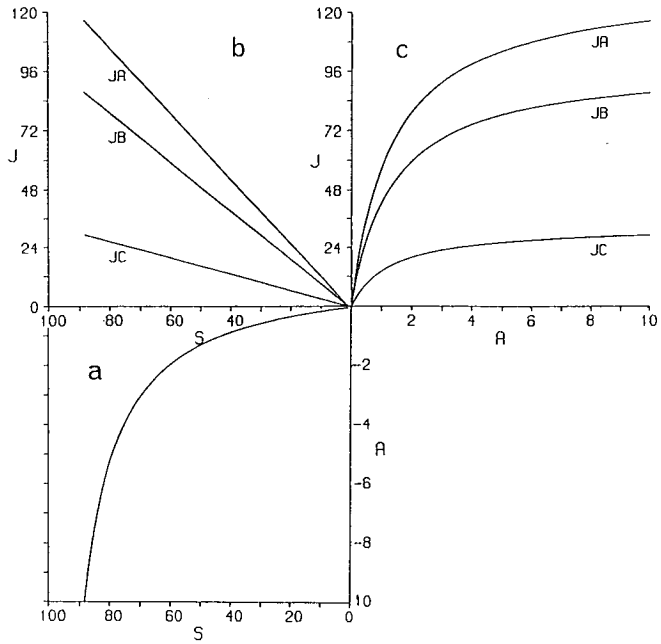


FIGURE 2.—Flux responses to changes in the enzyme activity of the common branch in a system with first order enzymes (all enzymes unsaturated). The effects are shown as a "reflection diagram" which 'decomposes' the effect into component functions of the final function. (*a*) How variation in the parameter A (group enzyme activity) changes the steady state value of the variable S (the branch point substrate). (see Figure 1). (*b*) Changes in S are now reflected in this panel which shows how changes in the variable affect the three fluxes dependent on it. This therefore shows how S is allocated from J_A to the two output fluxes J_B and J_C. (*c*) The flux changes are reflected back to A and show the resultant net effect of changes in A on the three fluxes. It will be noted that (*b*) shows linear relationships of S on J's. This implies that the Dominance Index (resulting from any three values of A) will be identical when measured in any of the three fluxes or in S. The relationships of the J's on A in (*c*) are simple hyperbolic functions. Any three values of A will give identical flux proportions in all three fluxes. It will also be noted that in (*b*) and (*c*) the fluxes are constrained by $J_A = J_B + J_C$.

The branched system with saturation

We now consider specifically the case of some degree of saturation in one of the branches. Let this be an enzyme E_j^B in the "B" branch. The system is represented in Figure 3.

The "A" and "C" branches are identical in structure to the nonsaturable system (Figure 1b). The saturable branch can be divided into three parts: (a) The "P" section of nonsaturable enzymes *proximal* to E_j^B. (b) The saturable enzyme E_j^B. E_j^B is saturable by either its product S_j or its substrate S_i or both, depending on the values of the Michaelis constants, M_j and M_i, respectively (see Eq. 1.1 in Appendix 1). (c) The "D" section of nonsaturable enzymes *distal* to E_j^B.

A quadratic expression for J_B is obtainable in terms of all the internal and external parameters (Appendix 3). The expression for J_c is also a quadratic and is also given in Appendix 3. These equations are difficult to manipulate algebraically, and, being non-linear, the matrix treatment does not apply. Their characteristics with respect to enzyme variation are better understood by evaluating specific parameter sets. (See *e.g.*, HOFMEYER 1986.) Branched pathways have been investigated previously [KACSER (1983) in terms of control analysis; LAPORTE, WALSH and KOSHLAND (1984) in terms of allocation of the fluxes; and SAURO, SMALL and FELL (1987) in terms of the matrix method giving branch distribution control coefficients].

Since we are interested in the possible effects of saturation on the dominance relations in J_B and J_C, we need to examine cases of 'high' and 'low' saturation. The degree of saturation may be described quantitatively by a simple saturation index, SAT:

$$\mathrm{SAT} = \frac{S_i/M_i + S_j/M_j}{1 + S_i/M_i + S_j/M_j}. \qquad (3)$$

SAT can take values as follows: (a) $S_i/M_i \ll 1$ and

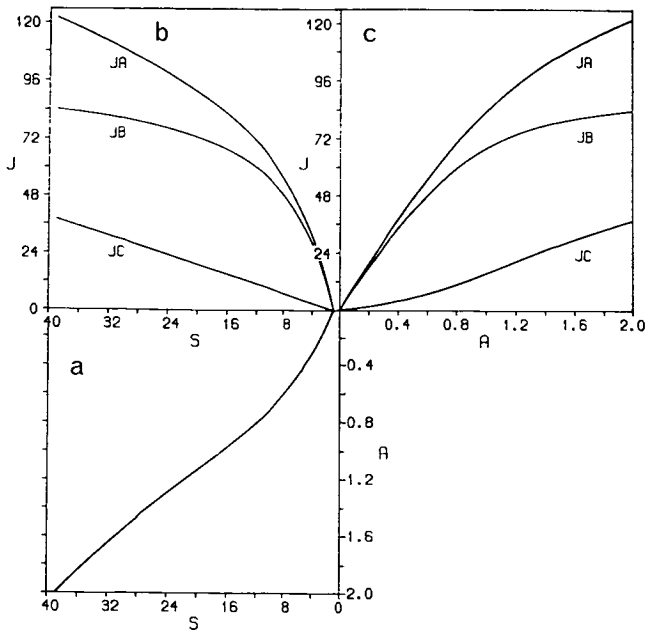


FIGURE 6.—Flux responses to changes in the enzyme activity of the common ("A") branch. Parameters as in Figure 4 with X_A at 100 (high saturation) and A varying from zero to 2. As explained in the legend to Figure 2, the effects are shown as a 'reflection diagram.' The "C" branch of first-order enzymes gives a linear flux response to S. The "B" branch, however, gives a damped response due to the presence of the saturable enzyme. This damping effect leads to the tendency for the "B" flux to give a 'more recessive' phenotype than the "C" flux (see Figure 5) and hence to different dominance indices.

of increasing saturation. Finally, Figure 6c shows the net effect of changes in A on the fluxes.

Thus, no matter which three values of S result from the three enzyme activities, the J_B phenotype will appear 'more recessive' than J_C. It is notable (Figures 5 and 6c) that in this case the flux in the nonsaturable branch, J_C, can give a Dominance Index greater than 0.5, *i.e.*, the mutant can therefore tend to be 'dominant' over the wild type (the heterozygote is nearer the mutant phenotype). This result, not previously observed, is due to the increasing slope of the J_C flux at low values of A (Figure 6c) in contrast to the monotonically declining change in J_B.

Enzyme variation in the nonsaturable branch: Figure 7 illustrates the dominance indices measured in the two fluxes for varying values of X_A where enzyme variation occurs in the nonsaturable branch. Here, we observe the opposite result from the effect of variation in the common branch. The flux measured through the nonsaturable pathway, J_C, is in this case a more recessive phenotype than that measured in the saturable branch. This result is also explainable in terms of the reaction of the fluxes to changes in the common substrate, S. This is illustrated in Figure 8. As in the case of variation in the common branch, J_B varies nonlinearly with S, higher values being damped due to the effect of saturation (Figure 6b). The flux in the

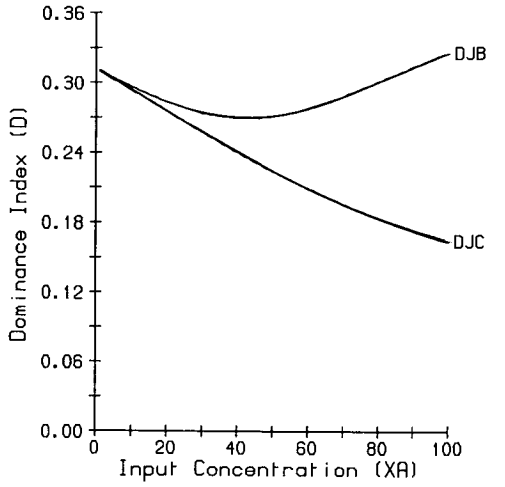


FIGURE 7.—The dominance indices of the two output fluxes with respect to enzyme variation occurring in the nonsaturable ("C") branch. Saturation is varied in the system by modulating the input X_A. The mutant enzyme activity is zero and the heterozygote activity is half wild type. The following parameters were used to generate the curves: $X_B = 1$; $X_C = 1$, $A = 1$; $P = 100$; $C = 10$ (wild type); $V_j = 50$; $M_i = 5$; $M_j = 1000$. All equilibrium constants were set to unity.

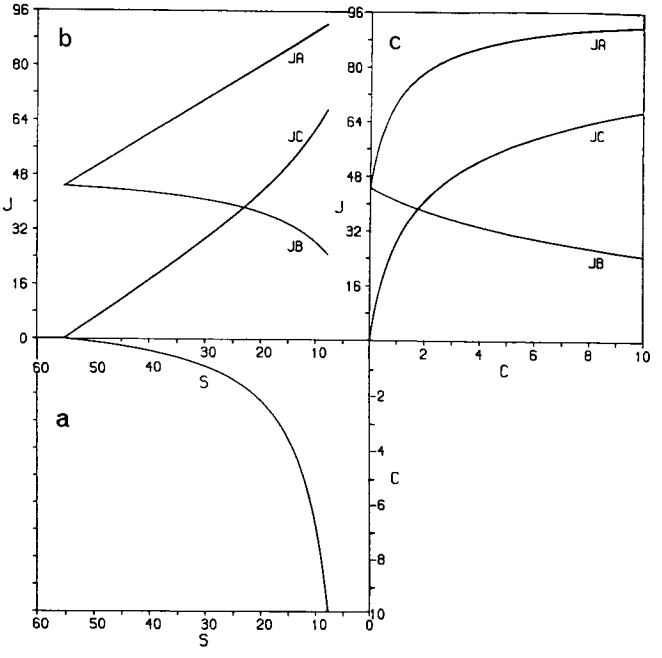


FIGURE 8.—Flux response to changes in enzyme activity in the competing "C" branch in the saturable system. This is a 'reflection diagram' (see Figure 2). Parameters as in Figure 7 with $X_A = 100$ and C varying from zero to 10. The "A" flux gives a linear response to changes in S, and "B" flux is damped due to the presence of saturation. The net result of these changes (c) is the tendency for J_C to give a more recessive phenotype than J_B.

common branch, J_A, is, however linear in S. Since $J_C = J_A - J_B$, J_C varies in this case nonlinearly with S, changes in which have been induced by enzyme variation in this pathway.

Enzyme variation in the saturable branch: In this case the numerical studies indicated, somewhat surprisingly, that the dominance indices measured in

This reaction is constrained by its stoichiometry and the rates of production of S_j and S_k must be identical. If two outputs arise from such stoichiometrically constrained fluxes, no question of competition for a shared metabolite arises. No difference in dominance relationship will therefore be observed, no matter what the saturation state or feedback conditions are.

For nonlinear systems we have had to use numerical simulations to assess the effects on independently variable fluxes. It has been shown that saturation at one enzyme can lead to differences in the dominance relations of two outputs. We have defined the directions of the differences in Dominance Indices for enzyme variation at different parts of the system. We have not explicitly considered systems with saturation in both the competing branches. The directions of any differences in dominance will depend on the relative saturation in each of the branches. The dominance relations in the presence of such nonlinear enzymes are dependent on their Michaelis constants as well as on the concentrations of external substances which affect the degree of saturation. Dominance and any possible differences are thus a function of the environment in which organisms are operating as well as of their genes.

The effect of saturation on the Dominance Indices is essentially through the resulting nonlinearity of some fluxes to substrate concentration. The presence of a feedback inhibition loop in the system can also lead to differences in the dominance between the outputs. The mechanism is essentially the same as for the case of saturation, with feedback inhibition introducing nonlinear relationships between fluxes and pools. This qualitative explanation of the behaviour of the saturable system and the system with feedback inhibition is robust to changes in parameters and consistent.

How do the above considerations relate to the dominance relations for pleiotropic effects of enzyme variation actually found *in vivo*? An enormous amount of work has gone into studies of human inborn errors of metabolism. These are frequently caused by near null-mutations at enzyme loci (HARRIS 1980). In most cases, heterozygotes are detectable if measurement of the enzyme activity is possible when it will show an activity which is usually the mean of wild type and mutant (STANBURY *et al.* 1983). In spite of this, the clinical phenotype of the heterozygote is 'recessive.' It has been shown that 'recessive' does not mean 'complete recessive' (*e.g.*, KNOX and MESSINGER 1958; BULFIELD and KACSER 1974; KACSER and BURNS 1981) as the heterozygote phenotype will invariably exhibit a small average difference from the wild type. Detection of heterozygotes will therefore depend on the ability to measure small differences and on the noise in the system. We have not, however, found any case in the literature where the heterozygote for a clinical phenotype is distinguishable from the wild type in the 'main effect' or any other manifestation of autosomal genes.

In Drosophila, a vast number of mutants are now known, many of which are null-mutants at enzyme loci. In these cases the situation appears to be the same as that for human inborn errors of metabolism—the mutations tend to be recessive for the 'main effect' and all the other pleiotropic effects (LINDSLEY and GRELL 1968).

There are, however, some cases in the literature where mutants show unequal dominance relations for different characters. In mice for example homozygotes for W^f alleles at the W locus are extensively depigmented, and there is no obvious pattern to the depigmentation (GUENET *et al.* 1979). Heterozygotes resemble the wild type except, however, for the presence of white spots on the forehead and belly. In these areas, the mutant allele is therefore 'dominant' for hair color, and in the rest of the animal, it is 'recessive.' The 'internal environment' of the gene in the different tissues must clearly be different. The two characters "pigment on dorsum" and "pigment on belly" can be compared to the example discussed in Figures 5 and 6 where the two branches would represent the melanin production in the two tissues. In the model X_A and X_B would be the same substance (melanin) in the two tissues. The enzymes would be the 'same' in the sense that they are specified by the same genes. The two tissues could however be different by, *e.g.*, sustaining different substrate concentrations and/or different activations or inductions of some of the enzymes. The effect of genetic substitution at one locus (W^f) could therefore have different consequences when the Dominance Index is measured in the two tissues.

A further example of differences in dominance relations occurs in the well known gene in pigs for halothane sensitivity (reviewed by WEBB 1981). Homozygotes for this allele are sensitive to the anaesthetic halothane, are stress susceptible, have improved meat colour and have improved performance for a number of economically important traits. Heterozygotes are not detectably different from the wild type for sensitivity to the anaesthetic, stress susceptibility, or meat colour, but are nearly intermediate for the other traits (*e.g.*, growth rate and carcass quality). The allele is therefore 'recessive' for some of the characters it affects, but is 'additive' for others. We have found one less clear cut example of this type of phenomenon in *Drosophila melanogaster*. The allele *scabrous-like* found in an abdominal bristle number selection line (HOLLINGDALE 1971) is a recessive semilethal, but has a substantial effect on bristles in the heterozygote. The biochemical bases of all these ef-

ysis. Additional relationships between elasticities and control coefficients. Eur. J. Biochem. **148:** 555–561.

FISHER, R. A., 1928 The possible modification of the response of the wild type to recurrent mutations. Am. Nat. **62:** 115–126.

FLINT, H. J., D. J. PORTEOUS and H. KACSER, 1980 Control of the flux in the arginine pathway of *Neurospora crassa*. Biochem. J. **190:** 1–15.

FLINT, H. J., R. W. TATESON, I. B. BARTHELMESS, D. J. PORTEOUS, W. D. DONACHIE and H. KACSER, 1981 Control of the flux in the arginine pathway of *Neurospora crassa*. Biochem. J. **200:** 231–246.

GRAHAM, A., 1981 *Kronecker Products and Matrix Calculus with Applications*. Ellis Horwood, Chicester.

GROEN, A. K., R. VAN DER MERWE, H. V. WESTERHOFF, R. J. A. WANDERS, T. P. M. ACKERBROOM and J. M. TAGER, 1982 Control of metabolic fluxes. pp. 9–37. In: *Metabolic Compartmentation*, edited by H. FRIES. Academic Press, New York.

GUENET, J.-L., G. MARCHAL, G. MILON, P. TAMBOURIN and F. WENDLEY, 1979 Fertile dominant spotting in the house mouse. J. Hered. **70:** 9–12.

HARRIS, H., 1980 *The Principles of Human Biochemical Genetics*, Ed. 3. Elsevier, Amsterdam.

HESS, B., 1973 Organization of glycolysis: oscillatory and stationary control. Symp. Soc. Exp. Biol. **27:** 105–132.

HEINRICH, R. and T. A. RAPOPORT, 1974 A linear steady state treatment of enzyme chains. Eur. J. Biochem. **42:** 89–102.

HOFMEYER, J. H. S., 1986 Steady state modelling of metabolic pathways: A guide for the prospective simulator. Cabios Rev. **2:** 5–11.

HOFMEYER, J. H. S., H. KACSER and K. J. VAN DER MERWE, 1986 Metabolic control analysis of moiety-conserved cycles. Eur. J. Biochem. **155:** 631–641.

HOLLINGDALE, B., 1971 Analysis of some genes from abdominal bristle number selection lines in *Drosophila melanogaster*. Theor. Appl. Genet. **41:** 292–301.

KACSER, H., 1983 The control of enzyme systems in vivo: elasticity analysis of the steady state. Biochem. Soc. Trans. **11:** 35–40.

KACSER, H., 1987 Dominance not inevitable but very likely. J. Theor. Biol. **125:** 505–506.

KACSER, H. and J. W. PORTEOUS, 1987 Control of metabolism: what do we have to measure? Trends Biochem. Sci. **12:** 5–14.

KACSER, H. and J. A. BURNS, 1973 The control of flux. Symp. Soc. Exp. Biol. **27:** 65–104.

KACSER, H. and J. A. BURNS, 1981 The molecular basis of dominance. Genetics **97:** 639–666.

KNOX, W. E. and E. MESSINGER, 1958 The detection in the heterozygote of the metabolic effect of the recessive gene for phenylketonuria. Am. J. Hum. Genet. **10:** 53–60.

LAPORTE, D. C., K. WALSH and D. E. KOSHLAND, 1984 The branch point effect. J. Biol. Chem. **259:** 14068–14075.

LINDSLEY, D. L. and E. H. GRELL, 1968 Genetic variations of *Drosophila melanogaster*. Carnegie Inst. Wash. Publ. 627.

MIDDLETON, R. J. and H. KACSER, 1983 Enzyme variation, metabolic flux and fitness: alcohol dehydrogenase in *Drosophila melanogaster*. Genetics **105:** 633–650.

SALTER, M., R. G. KNOWLES and C. I. POGSON, 1986 Quantitation of the importance of individual steps in the control of aromatic amino acid metabolism. Biochem. J. **234:** 635–647.

SAURO, H. M., R. SMALL and D. FELL, 1987 Metabolic control and its analysis: extension to the theory and matrix method. Eur. J. Biochem. **165:** 215–221.

STANBURY, J. B., J. B. WYNGAARDEN, D. S. FREDRICKSON, J. L. GOLDSTEIN and M. S. BROWN, 1983 *The Metabolic Basis of Inherited Disease*. McGraw-Hill, New York.

STUART, F., D. J. PORTEOUS, H. J. FLINT and H. KACSER, 1986 Control of flux in the arginine pathway of *Neurospora crassa*: effects of co-ordinate changes of enzyme concentration. J. Gen. Microbiol. **132:** 1159–1166.

WEBB, A. J. 1981 The halothane sensitivity test. pp. 105–124. In *Proceedings of Symposium on Porcine Stress and Meat Quality*, Edited by T. FROYSTEIN, E. SLINDE and N. STANDAL. Norwegian Agricultural Food Research Society, As.

WOODROW, I. E., 1986 Control of the rate of photosynthetic carbon dioxide fixation. Biochim. Biophys. Acta. **851:** 181–192.

WRIGHT, S., 1934 Molecular and evolutionary theories of dominance. Am. Nat. **63:** 24–53.

Communicating editor: D. L. HARTL

APPENDIX 1

Here we derive a set of equations for the fluxes in the nonsaturable branched system (Figure 1b). Consider the following monomolecular step within one branch, catalyzed by a Michaelis-Menten enzyme, at steady state.

$$\cdots \rightleftharpoons S_i \overset{E_j}{\rightleftharpoons} S_j \rightleftharpoons \cdots$$

The rate, v_j, of the reaction is given by

$$v_j = \frac{(V_j/M_i)(S_i - S_j/K_{i,j})}{1 + S_i/M_i + S_j/M_j}, \tag{1.1}$$

(CLELAND 1963), where V_j is the maximal velocity (V_{max}); M_i and M_j are the Michaelis constants (K_m) for the forward and backward reactions respectively; S_i and S_j are the concentrations of the substrate and product and $K_{i,j}$ is the equilibrium constant for the step which is, of course, independent of enzyme activity. When $S_i \ll M_i$ and $S_j \ll M_j$ (absence of saturation), Eq. 1.1 reduces to

$$v_j = e_j\,(S_i - S_j/K_{i,j}), \tag{1.2}$$

where $e_j = V_j/M_i$, the genetically determined enzyme activity.

Since the three branches of the system are symmetrical in structure, we can take the "A-pathway" (common branch) as an example. The steady state flux is obtained when all the individual rates in the branch are equal to one another and hence equal to the branch flux, *i.e.*, $v_1 = v_2 = v_3 = \cdots J_A$. All the intermediate pools will have time-invariant values. J_A is given by the solution of a set of linear simultaneous equations of the same form as Eq. 1.2:

$$J_A = e_1^A(X_A - S_1^A/K_{A,1}^A),$$
$$J_A = e_2^A\,(S_1^A - S_2^A/K_{1,2}^A),$$
$$\cdots$$
$$\cdots$$
$$J_A = e_l^A(S_{l-1}^A - S/K_{l-1,s}^A).$$

Solving these equations gives:

$$J_A = \frac{X_A - S/K_A}{1/e_1^A + K_{A,1}^A/e_2^A + \cdots K_{A,l-1}^A/e_l^A}, \tag{1.3}$$

where $K_A = K_{A,1}^A \cdot K_{1,2}^A \cdot K_{2,3}^A \cdot \cdots K_{l-1,S}^A$, *i.e.*, the equilibrium constant, $K_{A,S}$ between X_A and S. Similarly the equilibrium constant between X_A and S_{l-1} is $K_{A,l-1}$, etc. (Figure 1b).

Expression (1.3) can be reexpressed by grouping the sum of the reciprocals of enzyme activities and the equilibrium constants in Eq. 1.3 into a 'group enzyme activity,' A:

$$J_A = A(X_A - S/K_A), \tag{1.4}$$

where

$$1/A = \sum_{i=1}^{l} K_{A,i-1}/e_i^A$$

and the first term contains $K_{A,0} = 1$.

saturation containing the relevant Michaelis constants.

$$J_B = \frac{V_j/M_i(S_i - S_j/K_{i,j})}{1 + S_i/M_i + S_j/M_j}. \tag{3.3}$$

The variable S terms can be eliminated from the above three equations and from (1.4) and (1.6) to yield a quadratic in J_B.

$$\begin{aligned} J_B^2[M_i/(DM_j) - K_P\,(1/P + 1/(AK_A + C))] \\ + J_B[V_j/(DK_{i,j}) + X_BM_i/(M_jK_D) \\ + V_jK_P(1/P + 1/(A/K_A + C)) + M_i \\ + K_P(AX_A + CX_C/K_C)/(A/K_A + C)] \\ + X_BV_j/K_{i,j}K_D - V_jK_P(AX_A + CX_C/K_C)/(AK_A + C) = 0. \end{aligned}$$

For an asymptotically stable steady state this has only one positive real root.

The expression for J_C is obtained by eliminating S from Eqs. 2, 1.4 and 1.6. We obtain

$$J_C = \frac{X_A - X_C/K_{AC} - J_B/A}{1/A + 1/CK_A}, \tag{3.5}$$

where K_{AC} is the equilibrium constant for the reaction $X_A \rightarrow X_C$. The root of Eq. 3.4 (if known) can then replace J_B in (3.5) giving an expression for J_C in terms of parameters only.

APPENDIX 4

We derive equations for the fluxes in the branched system with feedback inhibition. At steady state, the set of simultaneous equations defining the fluxes in the system, defined in Figure 10 is given by (2), (1.4) and (1.6) together with four equations for J_B:

$$J_B = P(S - S_i/K_P), \tag{4.1}$$

$$J_B = Q(S_j - S_k/K_Q), \tag{4.2}$$

$$J_B = R(S_k - X_B/K_R), \tag{4.3}$$

where the equilibrium constant (K) terms are as defined in the legend to Figure 10 and P, Q and R represent the 'group enzyme activities' of the "P," "Q" and "R" sections respectively. A simple expression for the rate of the reaction $S_i \rightarrow S_j$ catalyzed by E_j can be

$$J_B = \frac{(V_j/M_i)(S_i - S_j/K_{i,j})}{1 + S_k/K_I} \tag{4.4}$$

(CLELAND 1963), where S_k is the concentration of the metabolite which acts as an allosteric inhibitor, and K_I is the inhibition constant. The solution to the set is a quadratic as follows:

$$\begin{aligned} J_B^2/(R/K_I) + J_B[1 + X_B/(K_RK_I) + V_jK_P/(PM_i) \\ + V_j/(QM_iK_{i,j}) + V_j/(RM_iK_{i,j}K_Q) \\ + V_jK_P/(M_i(A/K_A + C))] + V_jX_B/(M_iK_{i,j}K_QK_R) \\ - V_jK_P(AX_A + CX_C/K_C)/(M_i(A/K_A + C)) = 0. \end{aligned}$$

The flux through the "C" pathway is given, as before, by Eq. 3.5.

Directional Selection and Variation in Finite Populations

Peter D. Keightley and William G. Hill

Institute of Animal Genetics, University of Edinburgh, West Mains Road, Edinburgh EH9 3JN, Scotland

Manuscript received April 12, 1987
Revised copy accepted July 16, 1987

ABSTRACT

Predictions are made of the equilibrium genetic variances and responses in a metric trait under the joint effects of directional selection, mutation and linkage in a finite population. The "infinitesimal model" is analyzed as the limiting case of many mutants of very small effect, otherwise Monte Carlo simulation is used. If the effects of mutant genes on the trait are symmetrically distributed and they are unlinked, the variance of mutant effects is not an important parameter. If the distribution is skewed, unless effects or the population size is small, the proportion of mutants that have increasing effect is the critical parameter. With linkage the distribution of genotypic values in the population becomes skewed downward and the equilibrium genetic variance and response are smaller as disequilibrium becomes important. Linkage effects are greater when the mutational variance is contributed by many genes of small effect than few of large effect, and are greater when the majority of mutants increase rather than decrease the trait because genes that are of large effect or are deleterious do not segregate for long. The most likely conditions for "Muller's ratchet" are investigated.

IN recent years there has been much interest in the production and maintenance of variation in populations by mutation, stimulated by the presence of abundant variation in natural and artificial populations at the protein and DNA levels. Also, the genome is now seen as a fluid entity with transposition a particularly potent force in generating molecular variability. Variation at the phenotypic level must also originate from mutation, but the rate at which such variation is generated has been thought to be slow. This belief was derived mainly from observations of experimental populations of Drosophila. For example, the gain from new mutations in bristle score variation is of the order of one thousandth of the environmental variation per generation (discussed by LANDE 1976; HILL 1982b), and mutagenesis experiments have failed to produce large amounts of new variation in such quantitative traits (CLAYTON and ROBERTSON 1964; KITAGAWA 1967; HOLLINGDALE and BARKER 1971).

Despite the apparent slowness of accumulation of new mutational variance, theoretical analyses of the interaction of mutation and natural selection in the absence of drift have shown that mutation may be a powerful force in maintaining variation in natural populations (LANDE 1976), although the extent predicted depends on assumptions in the model (TURELLI 1984). Theoretical studies in finite populations have concentrated on the combined effect of mutation and directional selection in influencing quantitative variability and selection response rates (HILL 1982a,b). The equilibrium variance of a quantitative character is attained more quickly in the presence of selection than in its absence, and is highly dependent on population size. Thus in the early generations of a selection experiment or breeding program the response from variation generated by new mutations is expected to be small. In later generations, however, the contribution to the total variation present and hence to the response can be very important, especially in large populations. The results of long-term selection experiments can be interpreted in light of these analyses. The continued response after 120 generations of directional selection for increased pupal weight in Tribolium (ENFIELD 1980), after at least 75 generations of selection for increased bristle score of Drosophila (YOO 1980), and after 76 generations for increased oil content in maize (DUDLEY 1977) were likely to have been strongly influenced by variation arising while the experiment was proceeding.

In a previous paper (KEIGHTLEY and HILL 1983) Monte Carlo simulation was used to investigate the effect of linkage on asymptotic selection responses in small populations with new mutations and it was concluded that the asymptotic response rate is little affected by linkage, especially for species with many chromosomes. Moreover, the variance of effects of mutants on the trait was not an important parameter of the model for asymptotic selection responses were little affected by whether the new mutational variance arose from a few genes of large effect or many genes of small effect. Mutant effects on the trait were, however, assumed to be symmetrically distributed about zero.

Selection leads to a reduction in the genetic variance between family means, which appears as a negative disequilibrium covariance component within families. With free recombination the within family variance component is simply given by $V_g/2$ because recombination completely eliminates disequilibrium within families, but only half of the total genetic variance is initially present within full-sib families.

The total additive variance in the population after one generation of selection is obtained by adding the between and within family components,

$$V_{A,t+1} = (1 - h_t^2 k)V_{A,t}/2 + V_{g,t}/2. \quad (4)$$

The recurrence relation (4) corresponds to Equation 9.30 of BULMER (1980). Its validity depends on a normal distribution of genotypic values in the progeny, since skew in the distribution can affect the amount of variation removed by selection, but the results of BULMER (1980, Ch. 9) and ZENG (1987) indicate that, in many cases, skewness effects are small and can be ignored. Here, we use simulation to investigate possible effects of such skewness.

Mutation: As with the genic variance, the additive variance increases by V_M units each generation from mutation. Equation 4 becomes

$$V_{A,t+1} = (1 - h_t^2 k)V_{A,t}/2 + V_{g,t}/2 + V_M. \quad (5)$$

Finite population size: With a poisson distribution of family size the expected reduction in the additive genetic variance in the population is by a proportion $1/2N$ in the absence of selection. With truncation selection, the within family variance is independent of the population size, but the expected reduction in the between family component is by a proportion $1/N$ due to sampling of parents with replacement. Equation 5 becomes

$$V_{A,t+1} = (1 - 1/N)(1 - h_t^2 k)V_{A,t}/2 + V_{g,t}/2 + V_M. \quad (6)$$

Linkage: Linkage does not affect the variance between family means after one generation of selection, but affects the within family variance by reducing the amount of variation recovered from the disequilibrium covariance component due to recombination between loci. If c_{ij} is the recombination fraction between loci i and j, the disequilibrium remaining in the within family component is given by

$$2 \sum_{i \neq j} \sum D_{ij,t} (1 - 2c_{ij}) a_i a_j = \Sigma D_t - 2\Sigma(D_t c),$$

where ΣD_t is defined by (1) and similarly $2\Sigma(D_t c) = \Sigma\Sigma_{i \neq j} D_{ij} a_i a_j c_{ij}$. More generally (6) becomes

$$V_{A,t+1} = (1 - 1/N)(1 - h_t^2 k)V_{A,t}/2 + [V_{g,t} + \Sigma D_t - 2\Sigma(D_t c)]/2 + V_M. \quad (7)$$

The recurrence relation for the disequilibrium component is obtained by combining equations (2) and (7) to give

$$\Sigma D_{t+1} = \Sigma D_t - \Sigma D_t/(2N) - \Sigma(D_t c) - h_t^2 k(1 - 1/N)(V_{g,t} - \Sigma D_t)/2. \quad (8)$$

Asymptotic variance

As $t \to \infty$ for finite N, the variances reach expected equilibrium values about which they fluctuate stochastically due to sampling. For free recombination (6) can be reexpressed as a quadratic

$$\tilde{V}_A^2(1 + k + 1/N - k/N) + \tilde{V}_A(V_E - 2NV_M) \cdot(1 + 1/N) - 2V_M V_E(N + 1) = 0, \quad (9)$$

where $\tilde{V}_A$ is the equilibrium value of V_A. Ignoring second order terms, (9) is approximated by

$$\tilde{V}_A^2(1 + k) + \tilde{V}_A(V_E - 2NV_M) - 2NV_M V_E = 0. \quad (10)$$

Thus, $\tilde{V}_A$ is a function of mutation rate and population size as their product NV_M.

For complete linkage ($c = 0$ for all possible pairs of loci) a quadratic in $\tilde{V}_A$ is obtained by combining (2) and (7),

$$\tilde{V}_A^2(Nk + 1 - k) + \tilde{V}_A(V_E - 2NV_M) - 2NV_M V_E = 0. \quad (11)$$

Here, the relationship of $\tilde{V}_A$ to V_M and N is not as simple as in (10), but is a function of NV_M and Nk.

When recombination is finite, the simultaneous recurrence relations (3), (7) and (8) do not appear to have a simple solution. Their properties were investigated by iterating until steady state was achieved with initial values of V_g, V_A and ΣD set at zero, as would be the case in an isogenic population. The effect of a finite chromosome length was modeled by dividing the chromosome into a large number of equivalent segments (typically 100) and calculating the recombination fraction and hence the disequilibrium contribution from each possible pair of segments. This method exactly models the case of infinitesimally small effects as the number of segments approaches infinity, but increasing the number of segments beyond 100 made almost no difference to the results. The total amount of recombination was specified by L, the map length of the chromosome, and HALDANE's (1919) mapping function was used to related recombination fraction to map distance (l_{ij}) between pairs of loci: $c_{ij} = [1 - \exp(-2l_{ij})]/2$. Previous analyses (AVERY and HILL 1979) indicate that other models relating recombination fraction to map length (e.g. with crossover interference) make little difference in this type of model.

THE SIMULATION MODEL

The model is similar to that described by KEIGHTLEY and HILL (1983), except here we have simulated

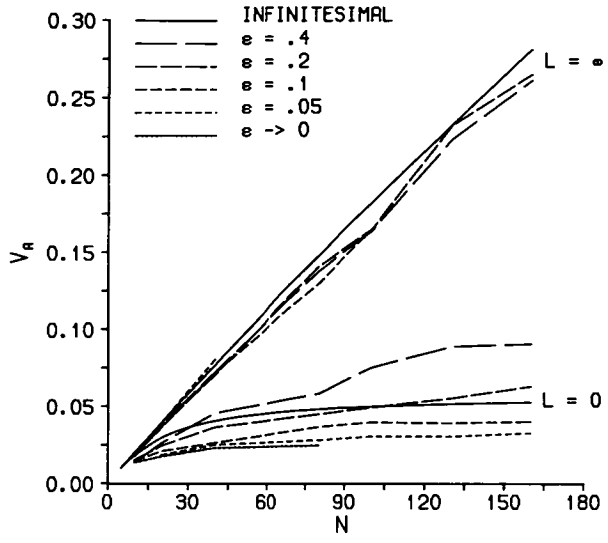


FIGURE 2.—The equilibrium variance is shown for the infinitesimal and simulation models for various population sizes and 50% truncation selection. $V_M/V_E = 10^{-3}$. In the simulation model, a range of values of ϵ and corresponding λ were used, with mutations coming from a symmetrical 'reflected double gamma' distribution ($P = 0.5$).

generations (200 for the populations simulated) were ignored. Thereafter the asymptotic response rate was calculated from the difference in mean value every other 10 generations and the mean genotypic variance, skew and kurtosis were computed every 10 generations. For a given V_M the computing time was approximately proportional to N^2 and inversely proportional to ϵ^2. So results for small ϵ (*e.g.*, 0.05) were only obtainable for N of 40 or less.

RESULTS

Comparison of the simulation and the infinitesimal model: Predictions of $\tilde{V}_A$ from both the Monte Carlo simulation and from the infinitesimal model for varying population size are shown in Figure 2. A value of V_M of $10^{-3}V_E$ was used, but in the simulation a range of sizes of effects was compared with corresponding values for the number of mutants per generation to satisfy (12).

With free recombination, the infinitesimal and simulation models are in good agreement. Surprisingly, the agreement is close even with relatively large effects and few mutants ($\epsilon = 0.4$). The disequilibrium present in the populations simulated can be estimated by subtracting the observed $\tilde{V}_A$ from the genic variance (given by $2NV_M$ in the infinitesimal case). As expected, with free recombination the amount of disequilibrium is small.

With complete linkage, the curves for different ϵ values differ substantially, larger values of ϵ giving higher predictions of $\tilde{V}_A$. The infinitesimal model is a poor predictor for complete linkage especially when effects are large, but it also overestimates $\tilde{V}_A$ when

TABLE 1

Equilibrium skewness of genotypic progeny values, computed as $g_1 = [\Sigma(X - \bar{X})^3/N]/\tilde{V}_A$, given for the case of no recombination ($L = 0$)

	Population size (N)			
	10	20	40	80
ϵ	Equilibrium skewness (g_1) among progeny			
→0	−0.0794	−0.147	−0.180	−0.184
0.05	−0.0592	−0.105	−0.149	−0.145
0.1	−0.0533	−0.0540	−0.0717	−0.0381
0.2	0.0142	0.0220	0.0465	0.0852
0.4	0.0289	0.0568	0.138	0.0735

$V_M/V_E = 10^{-3}$ and mutants come from a symmetrical 'reflected double gamma' distribution. Fifty percent truncation selection was simulated.

$\epsilon \to 0$. The overestimation can be explained by the presence of negative skew in the distribution of genotypic values of individuals (Table 1). Negative skew leads to a greater loss of variance each generation than predicted by the constant factor $i(i - x)$, and hence a lower $\tilde{V}_A$.

The effects of a finite amount of recombination are shown in Figure 3. The simulation and infinitesimal models agree at the free recombination limit but there is an increasing discrepancy at low recombination fractions. At the population sizes simulated, most of the effect of linkage is eliminated by one or two crossovers per chromosome per generation. The results are in agreement with those of KEIGHTLEY and HILL (1983) which used a fertility model of selection rather than the present viability model.

Asymmetrical distribution of mutational effects: Previous analyses (HILL 1982b; KEIGHTLEY and HILL 1983) have indicated that if the distribution of mutational effects is symmetrical (*i.e.*, the mutational variance contributed by negative and positive mutations is equal), then the shape of the density function of effects does not have much influence on selection responses and variation maintained.

Predictions of $\tilde{V}_A$ from simulations of different population sizes using the 'reflected double gamma' are plotted in Figure 4 for free recombination and values of P representing cases where mutants are mostly negative ($P = 0.1$), positive ($P = 0.9$) or symmetrically distributed ($P = 0.5$). The results show that $\tilde{V}_A$ is higher than the infinitesimal prediction (also shown in the figure) when mutants have predominantly positive effects and lower when most are negative. As the expected value of mutational effects approaches zero, however, the results approach the infinitesimal prediction. In the limit all the effects become infinitely small and the models must coincide. With finite effects, there are two reasons for the discrepancy from the infinitesimal prediction. Firstly, most negative mutations are lost almost immediately and contribute

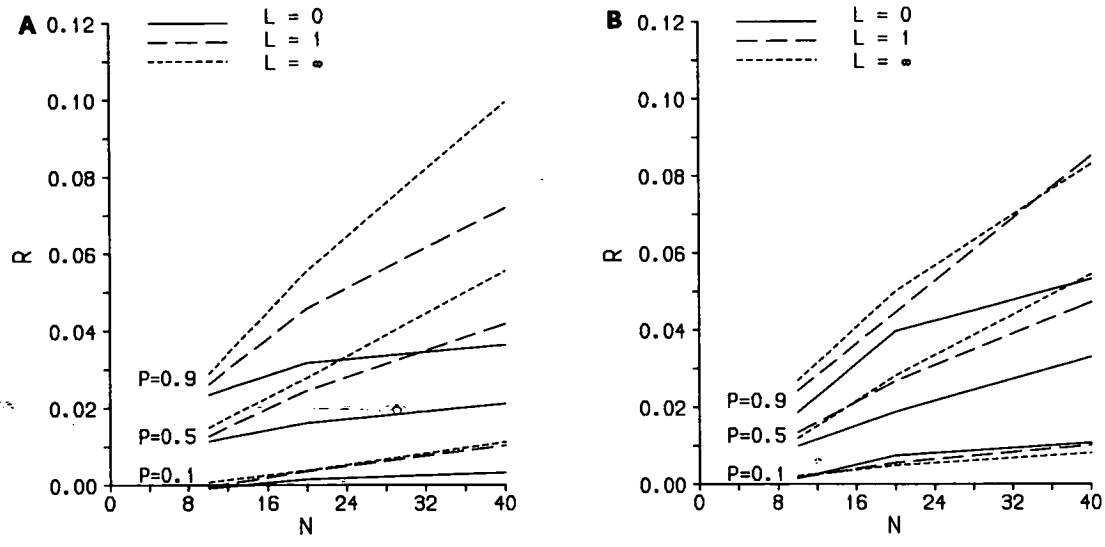


FIGURE 5.—Response rates where mutant effects have a reflected double gamma distribution with three values of P (proportion positive) in various population sizes. $V_M/V_E = 10^{-3}$ and 50% truncation selection. (A) $\epsilon = 0.1$; (B) $\epsilon = 0.4$.

values of P are plotted in Figure 5. Since Δ_m is independent of population size but $\tilde{V}_A$ is highly dependent, net responses become negative in small populations if most mutations are deleterious ($P = 0.1$).

Restating (12) and (14), $V_M \propto \lambda\epsilon^2$ and $\Delta_m \propto \lambda\epsilon$. It is clear that for a given V_M, as the expected magnitude of effects decreases ($\epsilon \rightarrow 0$) and hence the number of mutations increases, Δ_m must increase. Thus, if the new mutational variance is due to a large number of small negative effects, the mean value of a population will decline faster than if the mutational variance is due to a small number of large effects.

Paradoxically, the effects of linkage, *i.e.*, the difference between response rates for free recombination and complete linkage, are most severe when most mutations are positive ($P = 0.9$). Both positive and negative mutants interfere with each other's fixation probabilities (HILL and ROBERTSON 1966), but when most mutations are of positive value, there are more segregating so linkage effects are more important.

The response to selection from the fixation of freely recombining mutants can be approximated analytically if we assume that disequilibrium effects are small, so that the fate of each mutant is independent. In this case the response is given by

$$R = 2N\lambda \int_{-\infty}^{\infty} au(a)f(a)da \qquad (15)$$

where $f(a)$ is the density function of mutant effects and $u(a)$ is the fixation probability of mutants of effect a (HILL 1982a). For a gamma-distribution of mutational effects an approximation for R can be obtained by replacing $u(a)$ by the diffusion approximation of KIMURA (1957) for the fixation probability of additive genes [see HILL and RASBASH (1986) and Appendix].

Predicted response rates from simulation and Equation 15 are compared in Table 4. In general, the simulation agrees quite closely with the model of independent mutants. Comparing the results from the 'two-gamma' mutational distribution with those from the 'reflected double gamma' substantial differences in response rates can occur. The differences in response are consistent with the differences in variance maintained (Table 3). Also the differences are most extreme in small populations when the proportion of positive mutants (P) is 0.1. In this case, response rates are near zero so any difference is magnified.

DISCUSSION

Models: The computer simulation model is in itself of interest because it has been set up as far as possible in terms of known or measurable parameters, parental and progeny population sizes, map length of the chromosome, new mutational variance and distribution of mutational effects (assumed to be gamma form). The number of genes in the model is not fixed as in other models (*e.g.*, LANDE 1976; TURELLI 1984), but more naturally the number of loci with alleles segregating varies while the simulation is running. Furthermore, mutations which occur very close together on the chromosome can be considered either as alleles at separate loci or multiple alleles at the same locus. The model therefore connects and concurs simultaneously with the infinite locus models of BULMER (1971, 1976) and the 'infinite alleles' model of KIMURA (1965), and the possibility of intragenic recombination is accounted for. There are also similarities to a 'stepwise mutation' model (*e.g.*, TURELLI 1984). Any model of the mutational process, however, needs to be justified in terms of the effect on series of mutations on a gene (for say an enzyme) which in turn affects a quantitative character (say a flux), for which models have been

selection in a finite population. The model agree well with the Monte Carlo simulations both where effects are small ($\epsilon \rightarrow 0$), and also where mutational effects are relatively large. This behavior is consistent with the results of HILL (1982a) where independent genes give $\tilde{V}_A$ of $2NV_M$, irrespective of the mutational distribution. The higher fixation probability of mutants of large effect and their higher contribution to the variance in the character is nearly exactly balanced by their shorter fixation times and fewer number when compared to genes of small effect.

Asymmetry of mutant effects: When mutants come from an asymmetrical distribution the behavior is not as simple. As $\epsilon \rightarrow 0$ for any population size, the equilibrium additive variance will be essentially the same as predicted by the infinitesimal model. This will be true irrespective of the selection regime, selection only generating disequilibrium. With larger effects, $\tilde{V}_A$ becomes dependent on the proportion of mutants of positive effect (P). When the value of most mutational effects exceeds σ/Ni, $\tilde{V}_A$ is approximated by $4PNV_M$. This is so because the fixation probability of such mutants (and therefore the probability that they will reach intermediate frequencies and contribute substantially to the population variance) is proportional to a and independent of N. The number of mutants appearing in the population is, however, proportional to N. For small effects or in small populations, terms for which $|a| < \sigma/Ni$ become more important and $\tilde{V}_A$ will approach the infinitesimal prediction of $2NV_M$. Figure 4 shows, however, that effects must become very small or N very small before $\tilde{V}_A$ will be much different from $4PNV_M$.

With an asymmetrical mutational distribution, the distribution of genotypic values becomes skewed in the same direction as the mutational skew. A skewed distribution will lose more or less variance from directional selection depending on the sign of the skewness, so equilibrium variances are affected by such skewness. The simulations show that skewness is more important when mutational effects are large. At this point, we should mention that directional selection also generates skewness (positive) in the genotypic distribution (BULMER 1980, Ch. 9) so predicting the asymptotic distribution of genotypes when mutational effects are skewed is a difficult task.

The behavior of the system where genes are linked is also strongly affected by skewness in the genotypic distribution. When effects are small (*e.g.*, $\epsilon = 0.05$), the infinitesimal model overestimates the equilibrium genetic variance. The most likely explanation is a negatively skewed mutational distribution generated due to the loss of all but the best 'haplotype' and the presence of a 'tail' of individuals of lower value from mutation. This tendency to generate negative skewness is partially opposed by truncation selection generating positive skewness as mentioned earlier. The effect of linkage in generating skewness has been noted in earlier two locus studies (HILL and ROBERTSON 1966). Where effects are large, and therefore fewer mutations are occurring per generation, genes behave more as if they were independent and therefore higher V_A is maintained.

As a consequence of a negatively skewed mutational distribution, the rate of fixation of deleterious genes can exceed the rate for beneficial mutants and the population mean can decline; an effect corresponding to "Muller's ratchet" (MULLER 1964). We have identified a number of conditions necessary for the operation of the ratchet: (1) small population size since the fixation of deleterious mutants depends on chance; (2) many mutants of small effect (as opposed to a few of larger effect) since the 'mutation pressure' on the population mean is greater in this case; and (3) tight linkage since less standing variation will be available to oppose the mutation pressure. Linkage is also more important with small effects (*cf.* Figure 2).

Somewhat surprisingly the simulations show that linkage has most influence where most mutants are of positive value, and linkage effects can all but disappear when most are negative (*cf.* Figure 5). The explanation, however, is simple: deleterious genes almost never get fixed, while positive mutants get fixed with probability proportional (if independent) to a. In this latter case, however, linked positive mutants present simultaneously in the population can form unfavorable repulsion combinations leading to a reduction in fixation probabilities.

Directional *vs.* stabilizing selection: We should now point out that the free recombination results differ markedly from models of the maintenance of heritable variation in quantitative characters under mutation-stabilizing selection balance (*e.g.*, LANDE 1976; TURELLI 1984). In these models, the predicted equilibrium genetic variance $\tilde{V}_A$ is finite in an infinite population. With directional selection, $\tilde{V}_A$ will become infinite in an infinite population. The underlying cause of this discrepancy is the presence of mutants of positive effect on the character and hence on reproductive success with directional selection, but mutants of both positive and negative effect on the trait are deleterious for fitness with stabilizing selection (ROBERTSON 1956). Further consequences of the models are discussed elsewhere (HILL and KEIGHTLEY 1987).

We wish to thank the Agricultural and Food Research Council for financial support.

LITERATURE CITED

AVERY, P. J. & W. G. HILL, 1979 Variance in quantitative traits due to linked dominant genes in finite populations. Genetics **91:** 817–844.

BULMER, M. G., 1971 The effect of selection on genetic variability. Am. Nat. **105:** 201–211.

Genet. Res., Camb. (1988), **52**, *pp.* 33–43 *With 7 text-figures Printed in Great Britain*

Quantitative genetic variability maintained by mutation-stabilizing selection balance in finite populations

PETER D. KEIGHTLEY AND WILLIAM G. HILL
Department of Genetics, University of Edinburgh, West Mains Road, Edinburgh EH9 3JN

(*Received 25 October 1987 and in revised form 23 December 1987*)

Summary

Models of variability in quantitative traits maintained by a balance between mutation and stabilizing selection are investigated. The effects of mutant alleles are assumed to be additive and to be randomly sampled from a stationary distribution. With a two-allele model the equilibrium genetic variance in an infinite population is independent of the distribution of mutant effects, and dependent only on the total number of mutants appearing per generation. In a finite population, however, both the shape and standard deviation of the distribution of mutant effects are important. The equilibrium variance is lower when most of the mutational variance is contributed by few genes of large effect. Genes of small effect can eventually contribute substantially to the variance with increasing population size (N). The equilibrium variance can be higher in a finite than an infinite population since near-neutral alleles can drift to intermediate frequencies where selection is weakest. Linkage leads to a reduction in the maintained variance which is small unless linkage is very tight and selection is strong, but the reduction becomes greater with increasing N since more mutants segregate. A multi-allele model is simulated and it is concluded that the two-allele model gives a good approximation of its behaviour. It is argued that the total number of loci capable of influencing most quantitative traits is large, and that the distribution of mutant effects is highly leptokurtic with the effects of most mutants very small, and such mutants are important in contributing to the maintained variance since selection against them is slight. The weakness of the simple optimum model is discussed in relation to the likely consequences of pleiotropy.

1. Introduction

Many quantitative characters show considerable heritable variation in natural populations (Falconer, 1981; Mousseau & Roff, 1987; Roff & Mousseau, 1987). Explaining how such genotypic variation is maintained has been one of the most important and controversial problems of population genetics. The problem arises because of the widespread belief that stabilizing selection, in which the fittest individuals have values of the trait near some optimum, is ubiquitous in nature, but selection for an intermediate optimum is expected theoretically to deplete genetic variability (Robertson, 1956) and has been shown to do so experimentally (Gibson & Bradley, 1974; Kaufman, Enfield & Comstock, 1977). There is a certain irresistibility in arguments for the presence of an intermediate optima: for example, the date of egg laying in many northern passerine birds apparently has an optimum dependent on the availability of caterpillars for the young, which are only present for a brief period in early summer (Lack, 1968). Some of the most compelling evidence for selection for intermediate optima in natural populations comes from comparisons between sibling species of Drosophila, where parallel latitudinal clines for various traits have been shown to exist (David & Bocquet, 1975; Hyytia *et al.* 1985). The observation of an intermediate optimum at any single trait considered alone is, however, not in itself evidence of stabilizing selection as Robertson (1973) and others (Falconer, 1981; Rose, 1982; Hill & Keightley, 1988) have emphasized, because negative correlations between characters under directional selection can also generate such optima.

Genotypic mutations are the basic source of all heritable variation, but can a balance between mutation and selection alone explain the maintenance of observed high levels of genetic variation? This is an important question because variation in quantitative traits is believed to be the 'raw material' of evolution. Such variation also provides the basis for responses to

these occur independently. The increment in variance each generation from mutation is

$$V_M = \lambda E(a^2)/2 \quad (1)$$

(Hill, 1982*a*). Mutational effects are sampled from a time-invariant distribution. For modelling purposes the gamma distribution was chosen since it has a wide range of properties if suitable values are given to its two parameters. The density function of mutants having an increasing effect on the trait (illustrated in Fig. 1) is given by

$$f(a) = \alpha^\beta e^{-\alpha a} a^{\beta-1}/\Gamma(\beta) \quad (0 < a < \infty), \quad (2)$$

where $\Gamma(.)$ is the gamma function. The parameter α defines the scale of the distribution and β its shape. In practice, the scale was defined by the parameter $\epsilon = [E(a^2)/V_E]^{\frac{1}{2}} = [\beta(\beta+1)/\alpha^2]^{\frac{1}{2}}$, where V_E is the environmental variance. With shape parameter $\beta = 1$, $f(a)$ is an exponential distribution; as $\beta \to 0$ the distribution becomes increasingly leptokurtic with an increasingly large spike near $a = 0$ and a long tail; with $\beta \to \infty$, the distribution approaches the limiting case of all effects equal. The distribution is discussed in more detail by Hill & Rasbash (1986) (also see Kimura 1983, ch. 8). Mutants were assumed to have equal probability of increasing or decreasing the trait, with $f(-a)$ for $a < 0$ equalling $f(a)$ given by (2), i.e. a symmetric distribution over $-\infty < a < \infty$.

(*c*) *Selection.* The character is assumed to be under 'nor-optimal' stabilizing selection with the optimal phenotype fixed at zero. The phenotypic value of an individual is assumed to be the sum, X, of the contributions from each locus plus a random independent environmental effect of mean zero and variance $V_E = 1$. The relative fitness is given by

$$W(X) = \exp(-X^2/2w^2), \quad (3)$$

where w is a measure of the strength of stabilizing selection. Increasing w implies weaker stabilizing selection. With a multi-locus model where the population mean can vary due to gene frequency changes at any of the loci contributing to the character, Robertson (1956) showed that mutant alleles behave as under-dominant (i.e. the heterozygote is less fit than the homozygotes). The change of gene frequency at one locus under such stabilizing selection is given by

$$\Delta q = a^2(q-\tfrac{1}{2})\,q(1-q)/[4(w^2+\sigma^2)], \quad (4)$$

where σ^2 is the phenotypic variance (formally, the phenotypic variance less the genetic variance contributed by the locus). The term $w^2+\sigma^2$ is often called the strength of natural selection and referred to as V_s. This is equivalent to a model of heterozygote inferiority in fitness where $s = a^2/[8(w^2+\sigma^2)]$ is the fitness disadvantage of the heterozygote and there is a meta-stable equilibrium at $q = \frac{1}{2}$. Importantly, the strength of selection is proportional to the square of the allelic effect. Mutant genes are unconditionally deleterious and their selection is similar to that of genic selection with $\Delta q = s^* q(1-q)$, where

$$s^* = -a^2/[8(w^2+\sigma^2)]. \quad (5)$$

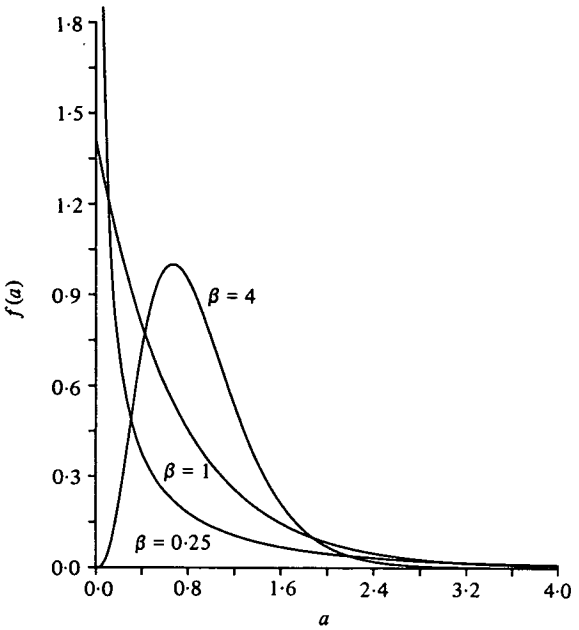


Fig. 1. Examples of the gamma distribution, $f(a) = \alpha^\beta e^{-\alpha a} a^{\beta-1}/\Gamma(\beta)$ for three values of the shape parameter β. The parameter α describes scale rather than shape and its value is such that $E(a^2) = 1$ for each curve.

3. Methods

(*a*) *Transition matrix.* With Robertson's (1956) result (equation 4) for the change of gene frequency it is possible to model the effects of continued stabilizing selection using a transition matrix. This method allows computation of the expected heterozygosity contributed by a new mutant during its lifetime in a population of N individuals assuming that no further mutation occurs at the locus while this mutant is segregating. The transition probabilities are defined by the square matrix **M** for the Wright–Fisher stochastic process with values

$$m_{jk} = \binom{2N}{k}(q+\Delta q)^k(1-q-\Delta q)^{2N-k} \quad (0 \leqslant j,\, k \leqslant 2N),$$

where $q = j/2N$ and Δq is given by (4). Let $\mathbf{f}^T(t)$ denote the row vector with elements $f_j(t)$ which are the probabilities of a population having gene frequency $j/2N$ $(0 \leqslant j \leqslant 2N)$ at the generation t. Thus for a new mutant, $f_1(0) = 1$ and all other elements are zero. The vector $\mathbf{f}^T(t)$ at generation t $(t > 0)$ is obtained from

$$\mathbf{f}^T(t) = \mathbf{f}^T(t-1)\,\mathbf{M}. \quad (6)$$

Let **h** denote a column vector whose elements are the expected heterozygosity at a locus with gene frequency $j/2N$ $(0 \leqslant j \leqslant 2N)$; so $h_j = 2j/(2N)(1-j/(2N))$. The expected cumulative heterozygosity, $H(a)$, contributed

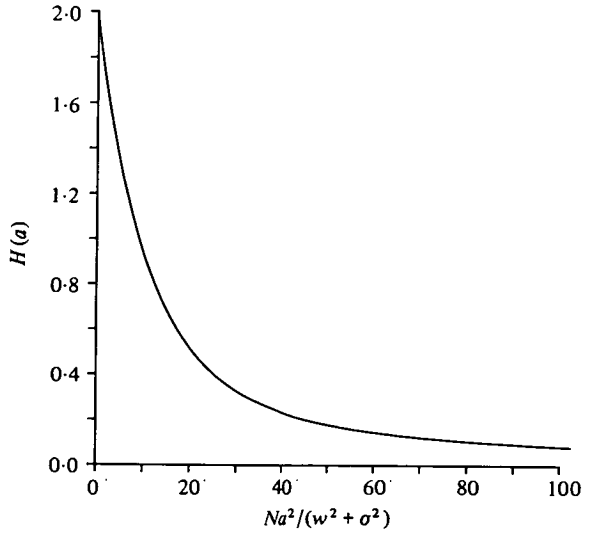


Fig. 2. Heterozygosity maintained during the lifetime of a mutant, $2\sum_{t=0}^{\infty} q_t(1-q_t)$, as a function of $Na^2/(w^2+\sigma^2)$, derived using the transition matrix.

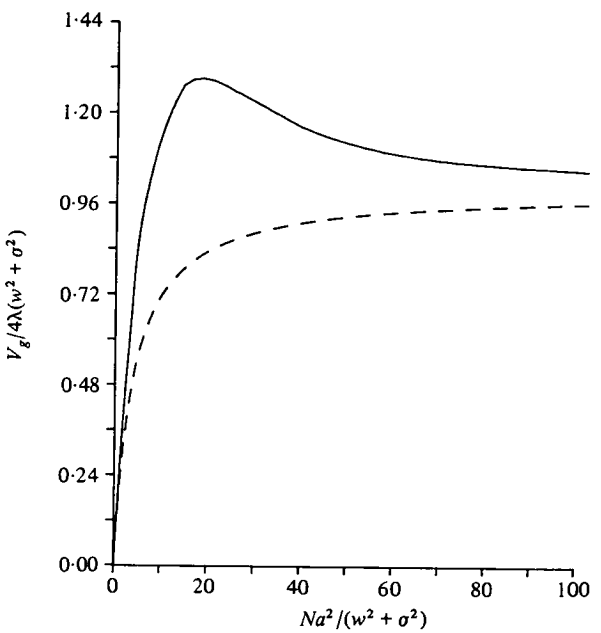


Fig. 3. Predicted equilibrium genetic variance, V_g, expressed as a proportion of that predicted for an infinite population, $4\lambda(w^2+\sigma^2)$, plotted against $Na^2/(w^2+\sigma^2)$. The curves compare predictions from the transition matrix (exact, ——) and from evaluation of equation (10) (approximate, – – –).

integrating (10) over a density function of mutant effects, but result (11) still holds. Equation (10) also shows that the equilibrium variance is (to first order) a function of $Na^2/(w^2+\sigma^2)$.

4. Results

(i) *Two-allele model*

(*a*) *Heterozygosity as a function of Ns.* The expected cumulative heterozygosity contributed by a mutant during its lifetime as a function of $Na^2/(w^2+\sigma^2) \propto Ns$ is illustrated in Fig. 2, computed using the transition matrix. $H(a)$ is bounded by the upper value of 2, where drift dominates, and the lower value of zero, where selection causes immediate elimination of the new mutant. The results in the following sections which give examples of V_g for different types of gamma distribution are all functions of the result in Fig. 2, and were generated by integrating numerically over this function with weighting according to the distribution of mutant effects (equation 7).

(*b*) *Variance maintained with genes of equal effect.* Fig. 3 shows the variance maintained as a proportion of that predicted in an infinite population (cf. equation (11)) as a function of $Na^2/(w^2+\sigma^2)$. The graph compares the results from the transition matrix and evaluation of (10). With increasing effects of drift ($Ns \to 0$), the variance maintained approaches zero; and as the effects of selection become more important (increasing N or $a^2/(w^2+\sigma^2)$), the relative variance maintained approaches the asymptote of 1. Interestingly, the results from the transition matrix indicate a maximum greater than the infinite population variance. If the effects of drift and selection are not too strong, the frequency of some mutants can approach the meta-stable point ($q = 0{\cdot}5$) where the expected change in gene frequency is zero. This possibility is not accounted for by equation (10), which assumes that the selection coefficient is constant and at its maximum. The presence of the maximum in Fig. 3 was confirmed by Monte Carlo simulation with equal mutant effects in a multi-locus model (results not given). Although not shown by Bulmer (1972), the maximum is also obtained by evaluating (9).

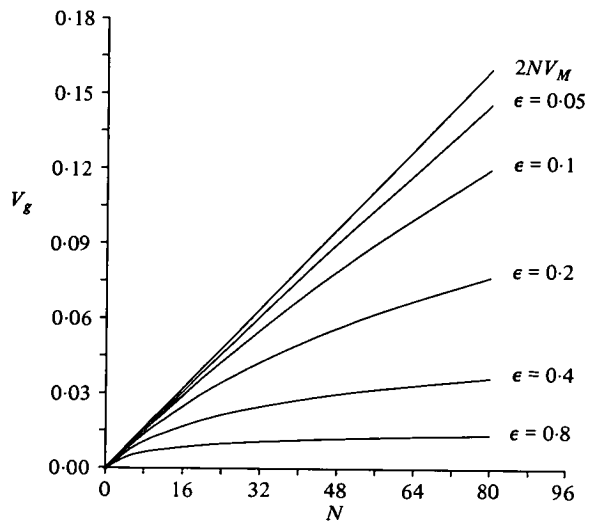


Fig. 4. Predicted equilibrium genetic variance maintained plotted against population size N. The curves were generated from the transition matrix, assumed a value of $V_M/V_E = 10^{-3}$, a gamma distribution of mutant effects, with $\beta = \frac{1}{2}$ and compare results for a range of $\epsilon = [E(a^2)/V_E]^{\frac{1}{2}}$ and corresponding mutation rate λ. The strength of stabilizing selection is given by $w^2 = \sigma^2 = 1$.

(*c*) *The influence of the scale of the mutational*

The models of Kimura (1965) and Lande (1976) (KL Gaussian models) assume that mutant effects differ only slightly from those already segregating, with the consequence that the steady-state distribution of allelic effects at the locus is normal. As Turelli (1984) pointed out, there is a fundamental discrepancy between the behaviour of the two types of model. Using Kimura's analysis, the equilibrium variance in an infinite population at a locus (V_{gL}) is

$$V_{gL} = [2V_{ML}(w^2+\sigma^2)]^{\frac{1}{2}}, \qquad (12)$$

where V_{ML} is the mutational variance input at the locus. This result can be derived by a different route. Assume at steady state a large number of alleles generates a normal distribution of allelic effects segregating at a locus. The variance maintained at the locus in a finite population can be obtained from the recurrence

$$V_{gL,t+1} = V_{gL,t}(1-1/2N_e)(1-V_{gL}k/2\sigma^2)+V_{ML} \qquad (13)$$

because the variance at the locus is reduced each generation by the factor $(1-1/2N_e)$ by drift and $(1-V_{gL}k/2\sigma^2)$ by selection, where k depends on the strength of selection and is the proportion of the phenotypic variance in the unselected individuals (Bulmer, 1980; Falconer, 1981, p. 180). With stabilizing selection and a normal distribution of phenotypic values, $k = \sigma^2/(w^2+\sigma^2)$. With infinite population size, (13) reduces to Kimura's (1965) formula (12) (ignoring second-order terms). Equation (13) also gives a solution for a finite population which is a quadratic in V_{gL}:

$$V_{gL}^2 k(2N_e-1)+2\sigma^2 V_{gL}-4N_e V_{ML}\sigma^2 = 0. \qquad (14)$$

This formula is similar to that obtained using the same assumptions by Latter (1970).

Fig. 7 compares the equilibrium genetic variance maintained for a range of population sizes for three different models using a gamma distribution of mutant effects with shape parameter $\beta = \frac{1}{2}$:

(i) Gaussian: the variance was computed from the solution to (14). This corresponds to the KL prediction.

(ii) 'Two allele': the variance was computed from the transition matrix and numerical integration, for a model of two alleles only per locus.

(iii) 'Multi-allele': the variance was computed by simulation of n discrete freely recombining loci with no intra-genic recombination, so the number of alleles which can segregate at any locus is not limited.

Also shown is the variance maintained by neutral genes which is simply $2NV_M$. The main points to note from this figure are: (*a*) all three models agree at small population sizes and mutant effects where drift is dominating; (*b*) the simulation of multiple alleles agrees with the KL Gaussian prediction only when mutant effects are small ($\epsilon = 0{\cdot}1$) and the mutation rate per locus is high ($\mu = 2\times10^{-3}$); (*c*) otherwise, with decreasing mutation rate, but correspondingly

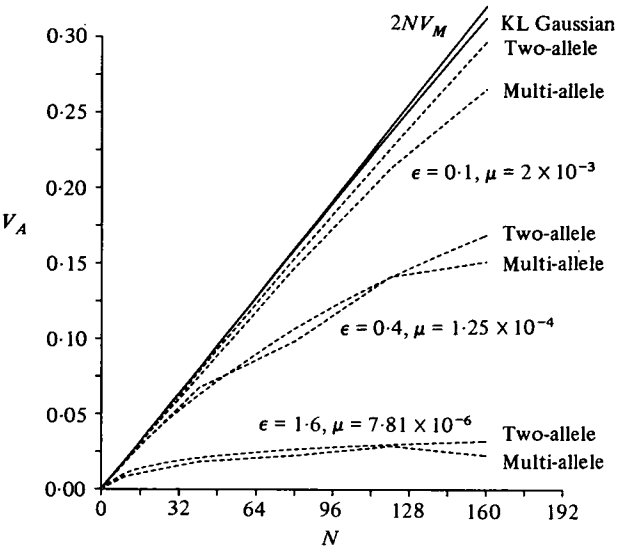


Fig. 7. Equilibrium additive variance plotted against population size, N, for three different models (see text). Also shown is the variance predicted for no selection, $2NV_M$. The simulation used 100 equally mutable freely recombining loci with mutation rates, μ, as shown, correponding values of ϵ so that $V_M = n\mu\epsilon^2/2 = 10^{-3}V_E$ and mutant effects sampled from a gamma distribution with $\beta = \frac{1}{2}$. The strength of stabilizing selection is given by $w^2 = 19$ $\sigma^2 = 19$. The standard errors of the results from the simulation are less than 5% of their mean.

increased magnitudes of mutant effects (e.g. $\epsilon = 1{\cdot}6$ and $\mu = 7{\cdot}81\times10^{-6}$), the simulation agrees better with the two-allele model. The simulation illustrates the difference between the KL approximation and Turelli's (1984) 'House of Cards' approximation. With the number of loci chosen for this example (100), the KL prediction hardly differs from neutrality. A larger number of loci would allow for a smaller standard deviation of the distribution of mutant effects for mutation rates per locus in line with those experimentally measured (Mukai & Cockerham, 1977). In this case, all three models would agree more closely at the population sizes shown, but as population sizes became much larger, would diverge as in Fig. 7 as the effect of selection becomes stronger relative to drift.

5. Discussion

(*a*) *Stabilizing selection and drift.* We have concentrated on a model where the mutation rate is sufficiently low or the population size sufficiently small that two alleles segregate at each locus. The consequences of such a model with infinite population size have been investigated previously (Latter, 1960; Bulmer, 1972) and an important conclusion was that the equilibrium genetic variance, V_g, is essentially independent of the effects of mutants on the trait, but depends only on the number of new mutants per generation. As a consequence of its independence of the effects of mutants, in an infinite population V_g is

Implications

The maintenance of genetic variation is a central problem in population biology, and the question of whether a mutation-stabilizing selection balance can maintain the observed levels of heritable variation has been frequently addressed (e.g. Lande, 1976, 1980; Turelli, 1984, 1985). The results show that, with finite population size, the shape of the mutational distribution has a strong influence on the genetic variance maintained under mutation-selection balance. There is little information concerning the shape of the distribution of mutational effects for any character, but can an informed guess be made from insights into biochemistry and molecular biology? In principle, all mutants, no matter where they occur in the genome, must have at least some effect on all characters, albeit very small. The interactive nature of metabolism, where the fluxes and metabolite pool concentrations, are systemic properties dependent to a greater or lesser extent on all enzymes in the 'metabolic map' (Kacser & Burns, 1973) tells us that there must be hundreds, if not thousands, of enzymes, variation in the activities of which will affect any character which is in some way controlled by the metabolism of the organism. Evidence for functional constraint in the genome (Kimura, 1983, ch. 7) at such sites as introns, silent (non-replacement) sites within coding sequences, and gene flanking sequences, suggests that there are many places in the genome capable of producing some small phenotypic effect. Thus it can be argued *a priori* that the distribution of mutant effects on complex quantitative characters is highly leptokurtic: most mutants are either of such trivial effect or so 'distant' from the character that they have almost no effect at all, but there is a smaller class of genes, more directly capable of influencing the trait with mutants of relatively large effect. The total number of mutants affecting a character is therefore high, much higher than an experiment designed to count polygenes would detect, but the effect of most of them is very small (see Robertson, 1967).

The difficulties in estimating the number and effects of mutants influencing a quantitative character are highlighted by the following illustrative example. Assume by genetic means only mutants showing an effect on the character of at least one-half of a standard deviation can be detected and the standard deviation of the mutational distribution is 1·6 units. If all effects were equal, then the genetic test would detect all the new mutants. If, however, the mutational distribution were more extreme (for example, gamma with shape parameter 0·25 (see Fig. 1)) then only 21 % of the new mutants would be detected but they would contribute most of the variance (96 %).

With this consideration in mind, estimates of the number of new mutants per generation affecting various quantitative traits in maize (Sprague, Russell & Penny, 1960; Russell, Sprague & Penny, 1963) seem rather high. These experiments gave estimates of λ for detectable mutants of about 0·06, implying, with, say, a mutation rate per locus of 10^{-5}, many thousands of loci at which mutations give sufficiently large effects to be detected. Such experiments, however, might now have to take into consideration the possibility of induction of 'mutator' genes (McClintock, 1950) in these crosses caused by movement of transposable elements known to be capable of affecting quantitative traits (Mackay, 1987). The rates of mutation may vary widely between populations as results of T. F. C. Mackay (personal communication) suggest, due to varying transposition rates.

(*a*) *Predicting maintained heritability – assigning values to parameters.* The important parameters are the mutational variance input per generation, the shape and scale of the mutational distribution, the strength and mode of operation of natural selection, and effective population size. As implied earlier, information is scarce on values of most of these parameters relevant to natural populations. If, however, it is assumed that $V_M/V_E = 10^{-3}$, the character is affected by a fairly extreme distribution of effects (i.e. gamma with $\beta = 0{\cdot}25$); most variance is contributed by mutants of fairly large effect (e.g. $\epsilon = 0{\cdot}4$ and therefore the mutation rate per genome $\lambda = 0{\cdot}0125$, implying *c.* 400 loci each mutating at *c.* $0{\cdot}3 \times 10^{-5}$); and a 'typical' value of $w^2 = 20\,\sigma^2$ (Turelli, 1984), then with $N_e = 10^3$, the maintained heritability would be about 21 %; with $N_e = 10^4$, the maintained heritability would be about 33 %, but there would be less than half of the genetic variance that would be present in an infinite population. It is notable that in an infinite population the maintained variance is proportional to $w^2 + \sigma^2$, and is very sensitive in a finite population to changes in $w^2 + \sigma^2$ over a wide range of parameters (flat part of curves in Fig. 5). If many characters are simultaneously subject to stabilizing selection, the value for w^2/σ^2 of 20 chosen in the above example may be smaller than typical (i.e. selection strength too strong) due to the genetic load that such selection would impose on the population. Thus, on the face of it, mutation-stabilizing selection balance is an attractive candidate for explaining the observed levels of heritable variation in populations that vary over a wide range in effective population size. The above calculations become less attractive when we consider the problems in estimating the strength of natural selection and in justifying the single character model of stabilizing selection. Such aspects have been discussed in detail by Turelli (1984, 1985).

(*b*) *Mode of action of natural selection.* The effect of pleiotropy is to reduce the genetic variation maintained since, for example, the selection coefficient against a mutant if stabilizing selection acts independently on each character, i, is proportional to $\sum[a_i^2/(w_i^2 + \sigma_i^2)]$. Clearly, the analysis could easily be

Robertson, A. (1967). The nature of quantitative genetic variation. In *Heritage from Mendel* (ed. R. A. Brink), pp. 265–280. Madison, Milwaukee and London: University of Wisconsin Press.

Robertson, A. (1973). The validity of the optimum model. *Advances in Applied Probability* **6**, 17–18.

Roff, D. A. & Mousseau, T. A. (1987). Quantitative genetics and fitness: Lessons from Drosophila. *Heredity* **58**, 103–118.

Rose, M. R. (1982). Antagonistic pleiotropy, dominance and genetic variation. *Heredity* **48**, 63–78.

Russell, W. A., Sprague, G. F. & Penny, L. H. (1963). Mutations affecting quantitative characters in long-time inbred lines of maize. *Crop Science* **3**, 175–178.

Slatkin, M. (1987). Heritable variation and heterozygosity under a balance between mutations and stabilizing selection. *Genetical Research* **50**, 53–62.

Sprague, G. F., Russell, W. A. & Penny, L. H. (1960). Mutations affecting quantitative traits in selfed progeny of doubled monoploid maize stocks. *Genetics* **45**, 855–866.

Turelli, M. (1984). Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. *Theoretical Population Biology* **25**, 138–193.

Turelli, M. (1985). Effects of pleiotropy on predictions concerning mutation-selection balance for polygenic traits. *Genetics* **111**, 165–195.