MOLECULAR BIOLOGY AND ANIMAL BREEDING (*)

Alan ROBERTSON

Institute of Animal Genetics
West Mains Road, Edinburgh, EH9 3JN, Scotland

SUMMARY

Developments in molecular biology in the last twenty years have radically altered accepted views on the nature of gene action and of protein synthesis and function. As yet they have had little direct impact on animal breeding. The author considers their implications for research in quantitative genetics and evolutionary studies in the light of recent developments in these fields.

INTRODUCTION

In this paper I want to deal with genetic variation as we see it from two aspects—firstly as quantitative geneticists and secondly, as workers in a field, which is expanding very rapidly at the present time, concerned with biochemical polymorphisms in animal populations. As quantitative geneticists we detect genetic variation in a particular measurement as a similarity between related individuals. If I was concerned with the genetics of hair colour I might say ,, identical twins usually have very similar hair colour '' and then, knowing that identical twins have identical genotypes, I would conclude that the variation in the population is to a considerable extent genetic in origin. In domestic animal populations, in which, with the exception of cattle, identical twins exist rarely or not at all, our measurements of this have to be made on individuals with a lesser degree of relationship such as parents and offspring, or groups of full or half sibs. Taking into account our knowledge from the basic diploid Mendelian mechanism of the extent to which relatives have genes in common, we can then make some estimate of the extent to which variations in any characteristic in the whole population are genetic in origin. This, the fundamental concept of quantitative genetics, is usually called the ,, Heritability '' of the measurement. The usefulness

of this concept essentially comes from the fact that, although different loci may contribute differently to these different phenomena, their relative proportional contribution is always the same. That is to say, if one locus contributes twice as much as another to the similarity between parents and offspring, it will also contribute twice as much to that between full brother and full sister, or half brother and half sister. The heritability then also allows us to make predictions of the probable responses of the population to selection. Note that, without knowing the number of loci we are working with and even without knowing with any precision the nature of gene action, we find ourselves able to describe adequately several different kinds of observations on the population and to predict how it will change with selection. However I must stress that a description in these terms is based on the effects of gene substitutions at the different loci and on the initial gene frequencies. In so far as we alter gene frequencies in a population by selection, we must as a necessary consequence expect to make changes in the genetic variation and in the heritability. Thus predictions of selection response are necessarily somewhat limited in time. As a matter of general experience of selection with laboratory and with domestic animals, such predictions are usually not useful for more than five generations and after that time we have to see to what extent the genetic variation has changed. Our theory then is essentially a statistical one and we can go quite a long way without ever penetrating to the level of individual loci.

BIOCHEMICAL POLYMORPHISMS

Now let us consider genetic variation from another point of view. The biochemists working with starch gels, or the blood group scientists working with immune reactions, can detect individual loci segregating in populations. The loci then are detected at a fundamental level, by examining their fairly immediate products. Our understanding of the basic relationship of the genetic material, the DNA, with its immediate products has increased tremendously in the past 15 years. For crude purposes we might then divide loci in terms of their immediate function into three different kinds.

i) Those whose primary function is to specify the sequence of amino acids in a particular peptide chain. The DNA sequence in the locus is firstly transcribed into a sequence in the messenger RNA which is then translated into a sequence of amino acids, with each three bases in the RNA sequence specifying a single amino acid. The synthesis of the peptide takes place on the ribosomes which are themselves made up of RNA and protein. These latter then are part of an all purpose machinery which can make any kind of protein.

ii) That part of the DNA which is concerned with the making of the machinery itself. Of this we do not know a great deal as yet although in some species the loci responsible for making the ribosomes have been identified.

iii) The third category of loci, which we must not leave out but of which we know almost nothing in higher organisms, are those concerned with regulation; they switch on or off the information genes and tell the machinery which particular

peptide chains to make. From the work of JACOB and MONOD in lower organisms like *Escherichia coli*, which led to the development of the concept of the operon, we have a fair idea of how regulation works in such simple organisms. Here some of the critical regulating substances appear to be peptide chains themselves. But in higher organisms we have little more than speculation (BRITTEN and DAVIDSON, 1969).

It would be idle to speculate which of these categories of loci may be most important in controlling quantitative variation but, because we know most about them, I am going to deal mostly with those loci which specify the sequence of amino acids in peptide chains. Such peptide chains may be the protein part of an enzyme molecule or of a molecule like hemoglobin or of a structural protein like casein or the muscle protein myosin. We can then ask various questions about the genetic variation in populations of animals at the loci specifying amino acid sequences. By the use of starch gel electrophoresis, we are in a position to answer the question as to what proportion of these loci show variability. Until recently, we had no clue as to what the answer was to this question because, as geneticists, we could only detect loci because they show variation. Classical genetics is a study of differences between individuals, and in that framework we cannot say at what proportion of loci there is variability. But if we accept present ideas of the relationship between the DNA sequence at the locus and the amino acid sequence in the peptide chain, there is a direct relationship between variation as seen on the starch gel and genetic variation at the locus producing that protein. If we see the same single band, or pattern of bands, in all individuals, then we can say that within the limits of this particular technique there is no variation at the locus producing this protein. The condition as to limits of the technique is an important one. It rests mainly on the charge differences between different variants of the same protein and not all amino acids substitutions which could arise by single base changes in the DNA will produce a change in electric charge. In fact it can be calculated that only about one third of such changes will do so. The two by now classical papers on this subject are by HARRIS (1966) on humans and by LEWONTIN and HUBBY (1966) in *Drosophila*. In both the approach was to take a variety of enzymes and detectable proteins and see in what proportion there was variation between individuals in the rate at which the protein migrated on the starch gel. HARRIS working with ten different systems, found variation in three and LEWONTIN and HUBBY found variation in eight out of 21. I may say that in my standard outbreeding population of *Drosophila melanogaster*, a different species from that used by LEWONTIN and HUBBY, I have found that of the ten enzyme systems that have been looked at there is variation in five. As, by this technique, we know that we must underestimate the proportion of loci which are varying, we can only come to the conclusion that in animal populations there is genetic variation at a vast number of loci which must run into the thousands. Although there are no critical estimates on other species, it remains obvious from the amount of variation which is continually being discovered at such loci that in animal populations there must be genetic variation at thousands of loci. We must presume that the original source of all this variation is mutation. As a single mutational step is presumably in the majority of cases a single base alteration in

the original DNA molecule, one might then expect that the consequence at the amino acid sequence level would be a single amino acid substitution. We have no evidence as to the nature of mutation in higher organisms but work on bacteria and viruses has in fact shown that simple mutational events lead to the replacement of individual amino acids. In human populations, very occasionally new haemoglobins, having arisen by mutation, can be detected because they are present in an individual and in neither of his parents. In these few cases the difference between the mutant form and the standard is a single amino acid change. In the peptide chain which has been most thoroughly analysed, that in the haemoglobin molecule, there are about 150 amino acids so that there are something of the order of 1,000 possible new sequences which could be produced from the standard sequence by a single amino acid substitution coherent with what we know of the triplet genetic code (LEHMANN and CARRELL, 1969). With this as a background we may then go on to ask the next question—how many different varieties of a given peptide chain are usually present within a population at any time. Again, we are somewhat limited by our technique of detecting differences in electric charge between molecules. The haemoglobin molecule can be used to illustrate the general situation. A fairly large number of variants are known which occur at very low frequencies within the population, presumably because they have arisen in the past few generations of a pedigree by mutation and are in the process of being selected out of the population again. In those populations in which we find true " polymorphisms " in which the variants are present at reasonable frequencies, we usually find only a comparatively small number of alternatives. In a sense the loci which do not show any variation can be included in this list as having only one sequence in the population. Some blood group workers may well object that they, in the well known B-system in cattle, have a system of several hundred alternative alleles. But the answer to this is that, on the basis of recent work, this must be considered as a complex locus controlled by many loci, in that there are now several examples of cross-overs within the system.

What is the nature of these polymorphic variants when we examine their amino acid sequence? In the case of the human haemoglobin polymorphisms, the two most frequent alternatives within a population have invariably been found to differ by a single amino acid. But more and more examples are beginning to be available of polymorphic variants which differ at two positions in the amino acid sequence, widely separated along the chain. They could not therefore have arisen from a single mutational event. The best known example of this is in the β-lactoglobulin polymorphism in cow's milk in which the two most frequent alternatives A and B, present in all breeds of cows looked at anywhere in the world, differ by two substitutions at positions 68 and 120 in the chain which is of total length 162 amino acids (FRANK and BRAUNITZER, 1967). Similar differences due to several amino acids have been found between the variants of the haemoglobins of sheep and of cattle. The significance of this is not clear. On the one hand it might suggest that these polymorphisms were of very long standing indeed so that multiple differences have had time to accumulate. On the other hand such a situation could equally well be found if the species had

split into two sub-species which had become fixed for different alternatives, followed by a subsequent reunification. It may be relevant that the only cases so far known of this are in domestic animals in which precisely such a process might have occurred during domestication.

Now we come to the really critical question. Why does all this genetic variation remain in the population — a question which exercises many population geneticists at the present time. We must first of all ask the question — how important is mutation in continually bringing new genetic variation into the population? Unfortunately we have almost no evidence as to what the mutation rate is for single base substitutions in higher organisms. Almost all the classical mutation work with *Drosophila* is based on lethal genes which may not be representative of the kind of mutations we are concerned with here but which in any case represent a whole class of mutations rather than a single individual. We can get some clues by looking at the frequency with which in human populations we find the comparatively rare variants which have presumably arisen in recent generations. From theoretical calculations it would seem likely that the average age of such variants in populations (which are never going to reach high frequencies and which will in the end be lost from the population) is between 10 and 20 generations. If we take published figures that in a survey of 8,000 Europeans, 5 such haemoglobins were found, of which there was no evidence of even a slightly endemic pattern in the population (LEHMANN and CARRELL, 1969), it can be calculated that the mutation rate to new detectable variants at the loci specifying the haemoglobin chains would be of the order of 10^{-6} per generation. Bearing in mind that this overall rate will be made up of 1,000 different amino acids ubstitutions but that only one third of these will be detected, we arrive at a final mutation rate for specific changes of the order of 3×10^{-9} per generation or once in 300,000,000 individuals. This is a very slow rate when one bears in mind that the total number of people who have ever lived in the British Isles is probably not more than 300,000,000. Nevertheless, in evolutionary terms this is quite a high rate. To me the fact that many loci are monomorphic in the sense that variants are very rarely found and that the majority of them which are polymorphic do not have very many alternatives is a strong argument for the suggestion that most mutations which cccur are probably somewhat disadvantageous and do not stay very long in the population.

The problem of how all this variation is maintained is exercising the minds of many geneticists at the moment. In some cases there is a markedly higher survival of heterozygotes such as with sickle-cell hemoglobin in humans in places where the incidence of malaria is high. But such large effects on the heterozygotes could not be tolerated at many loci. The homozygotes at the sickle cell locus are *inferior* and survive less well. As any individual would be bound to be homozygous at many of the segregating loci, if the effects at each locus were large the average survival of the population would be small. In my view, the most probable solution lies in heterozygote superiority with small effects at each locus, though not everybody in this field would agree with my views.

THE LOCI CONTROLLING QUANTITATIVE VARIATION

Now let me turn back to the loci concerned with quantitative variation. I said earlier that in our usual manipulations of selection and of heritability estimates we have no idea how many loci are affecting the character concerned. Quite a bit of recent experimentation has been concerned with this particular problem—attempting to identify individual loci controlling the genetic variation in characters like the number of bristles in a particular position in *Drosophila*. This is possibly a not very representative character but in working with *Drosophila* we can make use of techniques special to this organism to simplify our experimental work. However, I would think it quite possible that similar work could be done in species with only a small number of effective chromosomes and a fairly large number of marker genes available like the chicken. We may then ask the question — how many loci affect this character? I take the view that this is a stupid question — this character (like the majority of other characters) is the product of a long developmental process and as such will probably be affected by a great many loci, some of them to a rather small extent. Can we ask the question sensibly then? I believe that if I take my two extreme selected lines and ask how many loci I would need to implicate in explaining say 80 % of the selection response (leaving the remainder for genes with rather small effects) I think that as a result of THO-DAY's work in Cambridge (THODAY et al. 1964), and some of the work in my laboratory, we might answer ,, probably less than 10 ''. There has also been work on plants from ALLARD in California (WEHRHAHN and ALLARD, 1965) and from LAW (1967) in Cambridge working on wheat in which they detected quite large differences between different strains which they could associate with individual loci on particular chromosomes.

If one takes these populations and selects them, the contribution of a particular locus to immediate genetic improvement will be proportional to its contribution to the additive genetic variance. In highly selected populations then we may expect to have fixed the loci at which the effects are large, leaving us with a large number of loci still segregating at which the effects are small. I myself have been much interested in the possibility that in selection lines we may not succeed in fixing some genes which would be very useful to us. To do this I have to consider the effect of the population size on the response to selection, a branch of theory which has only recently been worked out (ROBERTSON, 1960). It turns out not very surprisingly that the chance of fixing a particular gene depends on its initial gene frequency as well as its effect on the character, so that one may miss a desirable gene either because it is rare or because its effect on the character under selection is small. From studies of the *Drosophila* lines in which the selection has not been very intense, I get the impression that the frequency of alleles leading to final response has been round about 0.5 in my initial population, but if I try very hard to pull out all the useful variation in my base population I begin to get hold of those variants which have a fairly large effect but which I had initially missed because the gene frequency had been small.

In selection theory it has been the habit to ignore the problem of linkage

between different loci although we all knew perfectly well that the genes we were manipulating were organised on chromosomes. With this work of locating individual loci the fact of linkage comes rather more close to us. In the last few years, aided by computer simulation work, it has been possible to get the problem of linkage in artificial selection somewhat more into focus (HILL and ROBERT-SON, 1966). The theoretical work then allows us to predict those conditions in which linkage is likely to be of importance in selection and in this the population size during selection appears as a very important parameter. It would seem likely from this work that linkage can in fact be ignored as a complication except in very special cases.

We are then in a position of having a great deal of information about the loci controlling biochemical variation on the one side and on the other, a statistical description of the genetic variation which only in very few cases can be split down into discussions of variation at particular loci. I myself would regard the investigation of the interrelationship between these two aspects of genetic variation as being one of the most important problems immediately before us.

BIOCHEMICAL POLYMORPHISMS AND QUANTITATIVE VARIATION

From the work on enzyme variation in man and in Drosophila and from the large number of such polymorphisms that we are now finding in domestic animal species, we must expect that there are many thousands of loci segregating in our animal populations. Those we have so far detected as such can be looked upon as the result of a fortunate accident — that we had available to us a biochemical or immunological technique which would detect the effect of that particular locus. Though we have detected many such, they must be only a small proportion of those remaining to be detected. If we then assume that there are only a small number of loci affecting a particular characteristic to an important extent, it is hardly surprising that the majority of investigations into the effect of blood group and other loci on quantitative characters have not been tremendously valuable. In some investigations, statistically significant effects have been discovered but they have none of them been of any practical significance with the possible exception of the B blood group locus in poultry in which some homozygotes are very definitely inferior in viability (BRILES and ALLEN, 1961). As a result, it has been claimed that the genetic constitution of grandparent lines has been deliberately altered to increase the frequency of given heterozygotes in the final commercial product.

How then should we proceed to locate those loci which are important in controlling the genetic variation in characters like milk yield or growth rate. If we select for either of these characters, what kind of biological processes are being altered in the animal? We have almost no clue as to how this question can be answered. Are particular cows high producing cows because they have some very efficient enzymes at critical points in the pathway of the synthetic processes or is it essentially a question of regulation so that a superior cow is merely one which diverts a high proportion of its food in the direction that we consider desi-

rable.   Recently a student in Edinburgh attempted to answer this question in reverse.   Several mutants are known in mice which cause a great increase in the amount of fat deposited.   The problem put was to identify the primary cause of action of one of these genes.   Previous work by biochemists had suggested that the prime cause was one of the enzymes in the Krebs carbohydrate cycle.   Howe-ver the final result of two years very hard work was that it is extremely difficult to answer this kind of question because of the buffering which is an essential part of any living process.   The buffering to reduce the effect of any disturbance on the end product usually means that the level of many intermediates in the process may be very much altered.   This was precisely what was found here — that the presence of the gene in homozygous form meant that the level of many enzymes were altered simultaneously and that it was impossible to locate the initial bio-chemical lesion.

Quite recently the two kinds of genetic variation were linked together by what can only be described as a fortunate coincidence.   For several years the minor milk protein, α-lactalbumin, has been known to be polymorphic in *Zebu* cattle, though not in European breeds, without anyone knowing what was the function of this protein.   First of all it was realised, from a study of the amino acid sequence of the protein, that it bore a striking resemblance to hen's egg white lysozyme whose function is to break down carbohydrate linkages in the cell wall of bacteria.   Then α-lactalbumin was discovered to be one of two proteins which together form lactose synthetase which is responsible for the joining of the two sugar molecules glucose and galactose to form lactose, the sugar constituent of milk (EBNER and BRODBECK, 1968).   So here we find the function of α-lactal-bumin is to join together two carbohydrate molecules whereas that of its analogue in the egg white is to break precisely such a linkage.

THE EVOLUTION OF PROTEINS

This surprising discovery of a similarity in basic sequence between two pro-teins in completely different species reminds me that I must not forget to mention perhaps what is the most fascinating genetic aspect of studies on the sequence of amino acids in proteins.   By sequencing a given peptide chain in different species, it is possible to follow the evolution of that particular chain by the gra-dual accumulation of amino acid differences.   For instance the α chain of human haemoglobin differs from that in the horse at 17 positions in the total sequence of 148.   The human α chain does not differ from that of the Gorilla whereas the two β chains differ by only one amino acid.   By means of this kind of information, an evolutionary tree can be built up solely from the present evidence of sequences and this tree agrees extremely well with that derived from the usual fossil evidence (ZUCKERKANDL and PAULING, 1965).   To give an example of the time scale, it is suggested that the average interval for the substitution of one amino acid for another in evolution is of the order of 10^7 years in this protein.

Further it becomes clear from sequence studies that the various peptide chains which go to make up the different haemoglobins must have a common evolu-

tionary source — in other words that there was a single locus concerned with production of a primitive haemoglobin-like molecule and that this had become duplicated in the course of evolution and that the related loci had now taken on rather different functions. More recently we have had evidence of the early stages of this process. Whereas the Sheep has apparently only one locus producing the α haemoglobin chain, the Goat and some of its relatives have two, producing two chains differing in sequence by 4 amino acids. The most interesting example is in the three loci controlling the production of casein in Cow's milk which is made up in the main of three fractions α_{s1}, β and $\varkappa$ (ASCHAFFENBURG, 1968). The amino-acid sequences of these proteins are now being analysed (RIBADEAU, DUMAS, GROSCLAUDE and MERCIER, 1970); from the gross amino-acid composition they are clearly closely related. Now genetic variation is known affecting each of these proteins separately. But when genetic analyses are carried out on animals heterozygous at two of the loci thee vidence shows quite clearly that the two locii must be very close together on the chromosome and in fact that all three locii form a triplet very closely linked together. Here we have a situation in which new loci have probably been produced by duplication of the original genetic material and in which the three still remain close together on the same chromosome. As an animal geneticist interested in evolutionary problems, I am delighted that the first indication of a triplet group of loci with a common origin, still very close together on the same chromosome, should come from studies of dairy cattle.

RÉSUMÉ

BIOLOGIE MOLÉCULAIRE ET SÉLECTION ANIMALE

La sélection basée sur la génétique quantitative procède par substitution génique à l'aveu-glette. Les fréquences géniques à l'intérieur d'une population se trouvent ainsi modifiées d'une manière inconnue du sélectionneur. De ce fait, la prédiction à long terme est limitée (5 générations au plus). Pour travailler avec plus de précisions, il faut étudier les gènes individuellement.

Depuis quinze ans nos connaissances sur toutes sortes de loci se sont accrues dans des pro-portions considérables, surtout sur la catégorie de gènes qui déterminent la séquence des acides aminés dans les chaînes peptidiques qui constituent la partie protéique des molécules d'enzyme ou de protéines structurales (caséine, myosine...).

D'après les études récentes on peut estimer que la moitié environ des loci de ce type présentent un di- ou un multi-allélisme. D'une manière générale, le multi-allélisme est limité, le locus de groupe sanguin B chez le bœuf restant une exception, en partie explicable par le fait qu'il s'agirait d'un locus complexe.

Le taux de mutation est très bas en valeur absolue bien qu'il soit assez élevé en terme d'évo-lution. Comme cependant les variants sont peu nombreux on peut penser que la plupart des mutants sont désavantageux pour l'espèce. Bien sûr il y a des effets spectaculaires d'overdomi-nance par exemple l'hémoglobine de l'anémie falciforme dans les zones paludéennes.

Un des problèmes concernant les loci contrôlant la variation quantitative est celui de leur nombre. Toutefois ainsi posée la question a peu de sens car tous les gènes peuvent affecter plus ou moins la réalisation d'un caractère. Si l'on se limite aux gènes qui contrôlent 80 p. 100 de la réforme à la sélection on peut souvent estimer que le nombre est inférieur à 10, dans quelques cas analysés.

Quant au linkage, de récentes expériences de simulation ont montré que, dans la majorité des cas point n'était besoin d'en tenir compte.

On peut penser que les gènes agissant sur le polymorphisme biochimique sont des milliers et que ceux qui sont connus actuellement le sont par suite du hasard. Comme on a vu qu'il y a seulement quelques loci agissant réellement sur un caractère donné, on ne doit pas s'étonner de ne pas encore les avoir trouvés. Pour le moment, il semble bien difficile de proposer une méthode

adéquate. De récentes études faites à Edimbourg sur des gènes qui chez la souris agissent sur le dépôt de graine n'ont pas été concluantes à cause des effets de tamponnement qui accompagnent les phénomènes vitaux. La présence d'un gène mutant sous une forme homozygote entraînait des altérations simultanées de plusieurs actions enzymatiques rendant impossible la localisation de la lésion biochimique initiale.

La surprenante découverte de la similarité entre chaînes protéiques diversement utilisées par des espèces aussi éloignées que la Poule et le Bœuf pose avec acuité le problème de l'évolution des protéines. En suivant une chaîne peptidique d'une espèce à l'autre, on peut espérer suivre l'évolution des espèces. Les observations montrent qu'il y a correspondances avec les déductions provenant de l'observation des fossiles. Quant à l'échelle des temps, on estime qu'il s'écoule en moyenne 10^7 années entre la substitution d'un acide aminé par un autre acide aminé dans un peptide donné.

Pour terminer, il faut évoquer le phénomène de duplication de loci provenant d'une source évolutive comme celui du locus contrôlant l'hémoglobine dans diverses espèces. Plus récemment même, le premier cas de triplication a été mis en évidence chez les bovins laitiers à propos des trois loci contrôlant les caséines α_{s1}, β et $\varkappa$ qui restent étroitement liés sur le même chromosome tout en subissant indépendamment des variations de structure biochimique.

REFERENCES

ASCHAFFENBURG R., 1968. Genetic variants of milk proteins: their breed distribution. *J. Dairy Res.*, **35**, 447-460.

BRILES W. E., ALLEN C. P., 1961. The blood group system of chickens II. *Genetics*, **46**, 1273-1293.

BRITTEN R. J., DAVIDSON E. H., 1969. Gene regulation for higher cells: a review. *Science*, **165**, 349-357.

DUMAS B., RIBADEAU, GROSCLAUDE F., MERCIER J. C., 1970. Structure primaire de la caséine β-bovine. *Eur. J. Biochem*, **14**, 451-459.

EBNER K. E., BRODBECK U., 1968. Biological role of α-lactalbumin: A review. *J. Dairy Sci.*, **51**, 317-322.

FRANK G., BRAUNITZER G., 1967. Primary structure of β-lactoglobulin. Hoppe-Seyler's Z. *Physiol. Chem.*, **348**, 1691.

HARRIS H., 1966. Enzyme polymorphisms in man. *Proc. Roy. Soc. B*, **164**, 298-310.

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

LAW C. N., 1967. The location of genetic factors controlling a number of quantitative characters in wheat. *Genetics*, **56**, 445-461.

LEHMANN H., CARRELL R. W., 1969. Variations in structure of human haemoglobin. *Brit. Med. Bull.*, **25**, 14-23.

LEWONTIN R. C., HUBBY J. L., 1966. A molecular approach to the study of genic heterogeneity in natural populations II. *Genetics*, **54**, 595-609.

ROBERTSON A., 1960. A theory of limits in artificial selection. *Proc. Roy. Soc., B*, **156**, 234-239.

THODAY J. M., GIBSON J. B., SPICKETT S. G., 1964. Regular responses to selection II. *Genet. Res.*, **5**, 1-19.

WEHRHAHN C., ALLARD R. W., 1965. The detection and measurement of the effects of individual genes involved in the inheritance of a quantitative character in wheat. *Genetics*, **51**, 109-119.

ZUCKERKANDL E., PAULING L., 1965. Divergence and Convergence in Proteins, in Bryson and Vogel *Evolving Genes and Proteins*, Academic Press, New-York.