Proc. Nat. Acad. Sci. USA
Vol. 71, No. 7, pp. 2848–2852, July 1974

On Some Principles Governing Molecular Evolution*

(population genetics/mutational pressure/negative selection/random drift)

MOTOO KIMURA† AND TOMOKO OHTA†

National Institute of Genetics, Mishima, Japan

Contributed by Motoo Kimura, May 1, 1974

ABSTRACT **The following five principles were deduced from the accumulated evidence on molecular evolution and theoretical considerations of the population dynamics of mutant substitutions: (*i*) for each protein, the rate of evolution in terms of amino acid substitutions is approximately constant/site per year for various lines, as long as the function and tertiary structure of the molecule remain essentially unaltered. (*ii*) Functionally less important molecules or parts of a molecule evolve (in terms of mutant substitutions) faster than more important ones. (*iii*) Those mutant substitutions that disrupt less the existing structure and function of a molecule (conservative substitutions) occur more frequently in evolution than more disruptive ones. (*iv*) Gene duplication must always precede the emergence of a gene having a new function. (*v*) Selective elimination of definitely deleterious mutants and random fixation of selectively neutral or very slightly deleterious mutants occur far more frequently in evolution than positive Darwinian selection of definitely advantageous mutants.**

Recent development of molecular genetics has added a new dimension to the studies of evolution. Its impact is comparable to that of Mendelism and cytogenetics in the past. Accumulated evidence suggests (1–8) that, as causes of evolutionary changes at the molecular (genic) level, mutational pressure and random gene frequency drift in Mendelian populations play a much more important role than the orthodox view of neo-Darwinism could lead us to believe.

In the present paper, we intend to enumerate some basic principles that have emerged from recent evolutionary studies of informational macromolecules. Of these, the first four are empirical, while the last one, which is theoretical, enables us to interpret the four empirical principles in a unified way.

(i) *For each protein, the rate of evolution in terms of amino acid substitutions is approximately constant per year per site for various lines, as long as the function and tertiary structure of the molecule remain essentially unaltered.* In their influential paper on the evolution of "informational macromolecules," Zuckerkandl and Pauling (9), noting that the mean evolutionary rates of globins are approximately equal per year among different lineages, suggested the existence of a molecular evolutionary clock. Actually, the idea of such a clock was implicit in the earlier writings of Ingram (10) and Jukes (11). The approximate constancy of the evolutionary rate in globins has since been confirmed by a number of authors (2, 12–14). For example, the number of observed amino acid differences between the α and β hemoglobin chains of man is approximately equal to that between the α chain of the carp and the β chain of man (12). Table 1 lists the numbers of amino acid sites in these two sets of comparisons that can be interpreted from the code table as due to a minimum of 0, 1, and 2 nucleotide substitutions. Also, the number of gaps due to insertion and/or deletion is listed. Since the human and carp α chains differ from each other at roughly 50% of the amino acid sites, the data suggest that the two structural genes coding for the α and β chains of hemoglobin have diverged independently of each other and to the same extent in the two lines since their origin by duplication which occurred possibly at the end of the Ordovician period. It is remarkable that mutant substitutions at gene loci coding for the α and β chains have occurred at practically the same average rates in the two separate lines that have evolved independently over nearly a half billion years. From these comparisons, the rate of amino acid substitution/site per year turns out to be about 0.9×10^{-9}. On the other hand, from comparisons of the α hemoglobin chains among various mammalian species, we obtain roughly the rate 10^{-9} site per year which is in good agreement with the above estimate. Although local fluctuations no doubt occur, constancy rather than variation of the evolutionary rate distinguishes the process of molecular evolution. This is

TABLE 1. *Comparison of amino acid differences between α and β hemoglobins*

Type of change*	Human α vs. human β	Carp α vs. human β
0	63	61
1	53	49
2	22	29
Gap	9	10
Total	147	149

* The numbers of amino acid sites that can be interpreted from the code table as due to a minimum of 0, 1, 2 nucleotide substitutions in two sets of comparisons involving the α and β hemoglobin chains. The number of gaps is also listed for each comparison.

* Contribution no. 1000 from the National Institute of Genetics, Mishima, Shizuoka-ken, 411, Japan.

† *We dedicate this paper to Dr. Hitoshi Kihara the former director of our institute in honor of his 80th birthday anniversary. He was really far-sighted when he wrote, as early as 1947, in relation to his outstanding cytogenetical work on the origin of cultivated wheat, "The history of the earth is recorded in the layers of its crust; the history of all organisms is inscribed in the chromosomes" (original in Japanese, ref. 42). With this paper we also celebrate the 25th anniversary of the National Institute of Genetics.*

particularly noteworthy since it is well-known (15) that there are enormous differences among evolutionary rates at the organism level; some forms have evolved very rapidly while others have stayed essentially unchanged over hundreds of millions of years (especially in organisms known as living fossils). Approximate constancy of the evolutionary rate per year has also been noted in cytochrome *c* and fibrinopeptides (4, 16, 17) although each has its characteristic rate; i.e., the evolutionary rate of cytochrome *c* is about $^1/_3$ while that of fibrinopeptides is roughly 4 to about 9 times that of hemoglobin. Constancy of the evolutionary rate per year has also been noted in albumin evolution of primates (18).

Recently, some authors have questioned the concept of a molecular clock by emphasizing local variation of evolutionary rates. For example, Goodman and his associates (19) emphasize that the evolutionary rates of the hemoglobin α chain slowed down in higher primates. Their method is based on estimating hidden mutant substitutions with the so-called "maximum parsimony" method, and accepts time spans from paleontological studies. In our opinion, the validity of their method (particularly the maximum parsimony principle) has to be tested in several cases rather than being taken for granted. More recently, Langley and Fitch (20) performed a somewhat more reliable analysis on the variation of evolutionary rates among the branches of a phylogenetic tree involving simultaneously the evolution of the α and β hemoglobins, cytochrome *c*, and fibrinopeptide A. They found that variation of evolutionary rates among branches ("legs") over proteins is significantly higher than expected by pure chance, with a χ^2 value about 2.5 times its degree of freedom. Since the expected value of χ^2 is equal to its degrees of freedom, their results mean that variation of evolutionary rate in terms of mutant substitutions among lines is about 2.5 times as large as that expected from chance fluctuations. Their estimation of the number of mutant substitutions is based on the assumption of minimum evolution, and it is likely that the estimation is biased in such a direction that lineages with more branches tend to show more hidden mutant substitutions. Yet, their results essentially agree with our previous analysis of the variation of evolutionary rates among lines using data on hemoglobins and cytochrome *c* (21). Namely, the observed variance of evolutionary rates among mammalian lines is roughly 1.5 to about 2.5 times the expected variance. However, the existing data indicate that, when averaged over a long period, the rate of evolution is remarkably uniform among different lineages, even though local fluctuations do occur.

We conclude, therefore, that constancy of evolutionary rate per year is valid as a first approximation. Such a constancy can be explained by the neutral mutation-random drift hypothesis if we assume that the rate of occurrence of neutral mutants is constant per year (4–6). Highly complicated and arbitrary sets of assumptions must be invoked regarding mutation, gene interaction, and ecological conditions as well as population size in order to explain the approximate constancy solely from the neo-Darwinian viewpoint. As predicted by one of us (12), it is likely that genes of "living fossils" in general have undergone essentially as many DNA base substitutions as corresponding genes in more rapidly evolving species. It is this constancy which makes the molecular data so useful and of such great potential value in constructing phylogenetic trees. Eventually, it will be possible to go far back into the history of life to clarify the early stage of evolution far beyond the capability of the traditional methods based on phenotypes.

(ii) *Functionally less important molecules or parts of a molecule evolve (in terms of mutant substitutions) faster than more important ones.* The rate of amino acid substitution has been estimated (with differing degrees of accuracy) for more than twenty different proteins as shown in Table 6-1 of Dayhoff (22). The highest rate is represented by fibrinopeptides (9×10^{-9}/amino acid per year according to their estimation) while the lowest rate is that of histone IV (0.006×10^{-9}). From this table it turns out that the median rate is 1.3×10^{-9}/amino acid per year. (represented by myoglobin in the table). This is not very different from 1.6×10^{-9}/amino acid per year which was estimated earlier by King and Jukes (2) as the average rate for seven proteins. Thus, hemoglobins show an evolutionary rate typical of those proteins that have been studied.

It is interesting to note that fibrinopeptides, the most rapidly evolving molecules, have little known function after they become separated from fibrinogen in the blood clot. The relationship between the functional importance (or more strictly, functional constraint) and the evolutionary rate has been beautifully explained by Dickerson (17) as follows. In fibrinopeptides, virtually any amino acid change (mutant substitution) that permits the peptides to be removed is "acceptable" to the species. Thus, the rate of evolutionary substitution of amino acids may be very near to the actual mutation rate. Hemoglobins, because they have a definite function of carrying oxygen and, so specifications for them are more restrictive than for fibrinopeptides, have a lower evolutionary rate. Cytochrome *c* interacts with cytochrome oxidase and reductase, both of which are much larger than it, and there is more functional constraint in cytochrome *c* than in hemoglobins. Thus cytochrome *c* has a lower evolutionary rate than hemoglobins. Histone IV binds to DNA in the nucleus, and is believed to control the expression of genetic information. It is quite probable that a protein so close to the genetic information storage system is highly specified with little evolutionary change over a billion years. Boyer *et al.* (23) reported that the δ chain of hemoglobin A_2 ($\alpha_2\delta_2$), which forms the minor component of adult hemoglobin, shows higher evolutionary rates and a higher level of polymorphism than the β chain which forms the major component A ($\alpha_2\beta_2$). This appears to agree with the present principle that less constraint enables more rapid change.

The evolutionary rate differs not only between different molecules but also between different parts of one molecule. For example, in both the α and β hemoglobin chains, the surface part of the molecule evolves nearly 10 times as fast as the functionally important heme pocket (7). In addition, two histidines binding to the heme are absolutely invariant throughout the entire history of vertebrate evolution extending nearly a half billion years (13). The Perutz model of hemoglobins (24) helps us greatly to interpret such observations in terms of structure and function of these molecules. More generally, if we consider the oil drop model of globular proteins (25), the inside of a molecule is filled with nonpolar (hydrophobic) amino acids, while the surface parts are occupied by polar (hydrophilic) amino acids. The functionally vital "active center" is located inside a crevice, and the rate of evolutionary substitutions of amino acids in this part is

expected to be very low. On the other hand, the surface parts are usually not very critical in maintaining the function or the tertiary structure, and the evolutionary rates in these parts are expected to be much higher. Another interesting example is the middle segment (C) of the proinsulin molecule. This part is removed when the active insulin is formed, and it is now known that this part evolves at the rate 4.4×10^{-9}/ amino acid per year, which is roughly 10 times as fast as that of insulin (6, 17). An additional example is afforded by the recent report of Barnard *et al.* (26). According to them, sequence 15–24 of pancreatic ribonucleases evolves at a very high rate comparable to rapidly evolving parts of fibrinopeptides, and this "hypervariability" can be correlated with a lack of any contribution of this part either to the enzymatic activity or to the maintenance of structure required for the activity. Incidentally, their Table 3 listing frequencies of amino acids in hypervariable segments suggests that in such regions there might still exist some selective constraint in amino acid substitutions, so that not all of the mutations are tolerated.

All the observations in this section allow a very simple interpretation from the neutral mutation-random drift hypothesis. Namely, in a molecule or a part of a molecule which is functionally less important, the chance of a mutant being selectively neutral (or very slightly deleterious) is higher, and therefore it has a higher chance of being fixed in the population by random drift. On the other hand, from the neo-Darwinian view-point, we must assume that a rapidly evolving part has an important functional role and is undergoing very rapid adaptive improvements by accumulating many advantageous mutations. It may be argued that the smaller the effect of a mutational change, the higher the chance of it being beneficial as Fisher (27) said, and therefore observations in this section can also be explained by positive natural selection. However, if the selective advantage of a mutant becomes small, then the chance of its fixation in the population becomes correspondingly small. Thus, apart from the problem of validity of Fisher's statement when applied to molecular data, it may not necessarily follow that the smaller the effect, the higher the rate of mutant substitution by natural selection.

(iii) *Those mutant substitutions that disrupt less the existing structure and function of a molecule (conservative substitutions) occur more frequently in evolution than more disruptive ones.* The conservative nature of amino acid substitutions was earlier noted by Zuckerkandl and Pauling (9). They also noted that the code table itself is conservative in that single base substitution often leads to both the substitution of a similar amino acid as well as a synonymous substitution. Since then, the conservative nature of substitutions has been amply documented in evolutionary studies of proteins (17, 22, 28, 29). Clarke (30) treated this problem in quantitative terms by using Sneath's (31) measure of chemical similarity of amino acids and by considering the regression of the relative frequency of evolutionary substitutions on the similarity. His results confirm the well-known fact that chemically similar substitutions occur more frequently than dissimilar ones.

The principle of conservative substitution holds also for nucleotide substitutions. In their extensive study on the evolution of transfer RNA, Holmquist *et al.* (32) found that, among the mispairings in the helical regions, G·U or U·G pairs that do not interfere with helicity occur much more frequently than other forms of mispairing; of 68 observed "non-Watson-Crick pairs," 43 turned out to be either G·U or U·G. For each transfer RNA molecule, the total number of mispairings in helical regions is limited to one or two, suggesting that beyond such a small number, a mutation leading to an additional mispairing becomes highly deleterious and rejected (it is likely that even the mutation causing the first mispairing is deleterious, but it can be fixed by random drift due to very small effect, see ref. 8); only when one of the existing mispairings is closed by a mutant substitution, is the molecule ready to accept a new mutation through random drift and/or selection. This offers an excellent model of Fitch's concept of concomitantly variable codons or "covarions" (33); according to him, only 10% of codons in cytochrome *c* can accept mutations at any moment in the course of evolution. He also found (34) that the proportion of covarions is about 35% in the hemoglobin α, but nearly 100% in fibrinopeptide A. A remarkable fact emerging from his analyses is that if the rate of amino acid substitution is calculated on the bases of covarions, cytochrome *c*, hemoglobin α, and fibrinopeptide A are all evolving at about the same rate. Fitch's covarion idea, we believe, has a clearer meaning now in the light of selective constraints involved in the secondary and tertiary structure necessary for the function of the molecule.

Similarly, one might expect that synonymous substitutions causing no change in amino acids would occur more frequently in evolution than nucleotide substitutions leading to amino acid change. From studies of amino acid sequences of tryptophan synthetase A-chains of three bacterial species, *Escherichia coli*, *Salmonella typhimurium*, and *Aerobacter aerogenes*, in conjunction with the estimated nucleotide sequence differences among the corresponding structural genes (determined by mRNA·DNA hybridization, Li *et al.* (35) obtained results suggesting that synonymous codon differences in the gene for tryptophan synthetase A chain are quite common. According to their estimate, there are about as many base differences that do not alter the amino acid sequences as those that alter the sequences. It is possible, as Li *et al.* point out, that not every synonymous substitution is completely neutral with respect to natural selection. Some of them might be subject to selective elimination based on structural requirement (such as the one involved in forming the secondary structure of the RNA molecule). However, because synonymous substitutions, in general, must have a higher chance of being selectively neutral or only very slightly deleterious (other things being equal) than mis-sense substitutions, they have a greater chance of becoming fixed in the population by random drift. One prediction that we could therefore make is that the slower the evolutionary rate of a protein molecule, the higher the ratio of synonymous to mis-sense substitutions.

(iv) *Gene duplication must always precede the emergence of a gene having a new function.* The importance of gene duplication in evolution has been noted earlier by the great Drosophila workers of the Morgan school (see ref. 5). The crucial point pertinent here is that the existence of two copies of the same gene enables one of the copies to accumulate mutations and to eventually emerge as a new gene, while another copy retains the old function required by the species for survival through the transitional period. Shielded by the normal counterpart in the corresponding site of the duplicated DNA segment, mutations that would have been rejected before

duplication can now accumulate, and through their accumulation, a stage is set for emergence of a new gene. The creative role which gene duplication plays in evolution has been much clarified by Ohno (36) in his stimulating book in which he considers new evidence based on modern molecular, cytological, and paleontological researches. Together with his recent paper (37), Ohno has made an important contribution to the modern evolutionary theory by bringing to light the remarkably conservative nature of mutant substitutions in evolution. Gene duplication, at the same time, must have caused a great deal of degeneration in duplicated DNA segments. This is because many mutations, which would have been definitely deleterious before duplication, become neutral or only very slightly deleterious after duplication, thus enabling them to spread in the population by random drift (38, 39).

(v) *Selective elimination of definitely deleterious mutants and random fixation of selectively neutral or very slightly deleterious mutants occur far more frequently in evolution than positive Darwinian selection of definitely advantageous mutants.* This is an extended form of the neutral mutation-random drift hypothesis, and is based on the thesis put forward by one of us (8) which argues that very slightly deleterious mutations as well as selectively neutral mutations play an important role in molecular evolution. Adaptive changes due to positive Darwinian selection no doubt occur at the molecular level, but we believe that definitely advantageous mutant substitutions are a minority when compared with a relatively large number of "non-Darwinian" type mutant substitutions, that is, fixations of mutant alleles in the population through the process of random drift of gene frequency. We emphasize that neutral or nearly neutral mutations should be considered not as a limit of selectively advantageous mutants but as a limit of deleterious mutants when the effect of mutation on fitness becomes small. In other words, mutational pressure causes evolutionary change whenever the negative-selection barrier is lifted. As an application of this principle, let us consider the evolutionary change of guinea pig insulin. Although the insulin (A and B segments) in general has a very low evolutionary rate (about 0.33×10^{-9}/amino acid per year), guinea pig insulin is exceptional in that it diverged very rapidly with the estimated rate of 5.3×10^{-9}/amino acid site per year (2). From the neo-Darwinian point of view, one might naturally consider such a rapid evolutionary change the result of adaptive change by natural selection. In fact, even King and Jukes (2) in their paper "Non-Darwinian Evolution" invoked "positive natural selection" to explain the rapid change. We suggest that guinea pig insulin lost its original selective constraint in the process of speciation. This allowed the accumulation of mutations which before would have been rejected. This inference is supported by a recent report of Blundell *et al.* (40) who studied the three-dimensional structure of insulin molecules. According to them, guinea pig insulin is accompanied by the loss of zinc in the islet cells (coinciding with the loss of usually invariant histidine B10). This suggests a drastic change in the tertiary structure. It is assumed then that, with the loss of the zinc constraint, mutations in guinea pig insulin started to accumulate at a very high rate approaching the rate in fibrinopeptides (the rate that might be called the fibrinopeptide limit).

When we consider the action of natural selection at the molecular level, we must keep in mind that higher order (i.e., secondary, tertiary, and quaternary) structures rather than the primary structure (i.e., amino acid sequence) are subject to selective constraint, usually in the form of negative selection, that is, elimination of functionally deleterious changes. The existence of selective constraint, often inferred from nonrandomness in amino acid or nucleotide sequences, does not contradict the neutral mutation-random drift hypothesis. Incidentally, it is interesting to note that the fibrinopeptide rate, when expressed in terms of nucleotide substitutions, is roughly equal to the rate of nucleotide substitution in the DNA of the mammalian genome (39). We note also that accumulation of very slightly deleterious mutants by random drift is essentially equivalent to the deterioration of environment, and definitely adaptive gene substitutions must occur from time to time to save the species from extinction.

Although clearly documented cases at the genic level are rather scarce, there is not a slightest doubt that the marvellous adaptations of all the living forms to their environments have been brought about by positive Darwinian selection. It is likely, however, that the ways in which mutations become advantageous are so opportunistic that no simple rules could be formulated to describe them. On the whole, mutations are disadvantageous, and, when a mutant is advantageous, it can be advantageous only under restricted conditions (41). We note also that difference in function at the molecular level, does not necessarily lead to effective natural selection at the level of individuals within a population.

In the past half century, with the rise of neo-Darwinism or more precisely, the synthetic theory of evolution, the claim that mutation is the main cause of evolution has completely been rejected. Instead, the orthodox view has been formed which maintains that the rate and direction of evolution are almost exclusively determined by positive natural selection. We believe that such a view has to be re-examined, particularly regarding evolutionary changes at the molecular level. We think that evolution by mutational pressure is a reality.

We thank Drs. T. H. Jukes and J. L. King for stimulating discussions which helped greatly to compose the manuscript. Especially, we are indebted to Dr. Jukes for critically reviewing the first draft and offering many suggestions for improvement. Thanks are also due to Drs. J. F. Crow and E. R. Dempster for reading the manuscript and offering suggestions for improving the presentation.

1. Kimura, M. (1968) *Nature* **217**, 624–626.
2. King, J. L. & Jukes, T. H. (1969) *Science* **164**, 788–798.
3. Crow, J. F. (1969) *Proc. XII Intern. Congr. Genetics (Tokyo)* **3**, 105–113.
4. Kimura, M. & Ohta, T. (1971) *J. Mol. Evolut.* **1**, 1–17.
5. Kimura, M. & Ohta, T. (1971) *Theoretical Aspects of Population Genetics* (Princeton University Press, Princeton, N.J.).
6. Kimura, M. & Ohta, T. (1972) *Proc. 6th Berkeley Symp. on Math. Stat. and Probability* **5**, 43–68.
7. Kimura, M. & Ohta, T. (1973) *Genetics (Sup.)* **73**, 19–35.
8. Ohta, T. (1973) *Nature* **246**, 96–98.
9. Zuckerkandl, E. & Pauling, L. (1965) in *Evolving Genes and Proteins*, eds. Bryson, V. & Vogel, H. J. (Academic Press, New York), pp. 97–166.
10. Ingram, V. M. (1961) *Nature* **189**, 704–708.
11. Jukes, T. H. (1963) *Advan. Biol. Med. Phys.* **9**, 1–41.
12. Kimura, M. (1969) *Proc. Nat. Acad. Sci. USA* **63**, 1181–1188.
13. Jukes, T. H. (1971) *J. Mol. Evolut.* **1**, 46–62.
14. Air, G. M., Thompson, E. O. P., Richardson, B. J. & Sharman, G. B. (1971) *Nature* **229**, 391–394.

15. Simpson, G. G. (1944) *Tempo and Mode in Evolution* (Columbia Univ. Press, New York).
16. Margoliash, E., Fitch, W. M. & Dickerson, R. E. (1968) *Brookhaven Symp. Biol.* **21**, 259–305.
17. Dickerson, R. E. (1971) *J. Mol. Evolut.* **1**, 26–45.
18. Sarich, V. M. & Wilson, A. C. (1967) *Proc. Nat. Acad. Sci. USA* **58**, 142–148.
19. Goodman, M., Barnabas, J., Matsuda, G. & Moore, G. W. (1971) *Nature* **233**, 604–613.
20. Langley, C. H. & Fitch, W. M. (1973) in *Genetic Structure of Populations* ed. Morton, N. E. (Univ. Press of Hawaii, Honolulu), pp. 246–262.
21. Ohta, T. & Kimura, M. (1971) *J. Mol. Evolut.* **1**, 18–25.
22. Dayhoff, M. O. (1972) *Atlas of Protein Sequence and Structure 1972* (National Biomedical Research Foundation, Silver Spring, Md.).
23. Boyer, S. H., Crosby, E. F., Thurmon, T. F., Noyes, A. N., Fuller, G. F., Leslie, S. E., Shepard, M. K. & Herndon, C. N. (1969) *Science* **166**, 1428–1431.
24. Perutz, M. F. & Lehman, H. (1968) *Nature* **219**, 902–909.
25. Dickerson, R. E. & Geis, I. (1969) *The Structure and Action of Proteins* (Harper & Row, New York, Evanston, London).
26. Barnard, E. A., Cohen, M. S., Gold, M. H. & Kim, Jae-Kyoung (1972) *Nature* **240**, 395–398.
27. Fisher, R. A. (1930) *The Genetical Theory of Natural Selection* (Clarendon Press, Oxford).
28. Epstein, C. J. (1967) *Nature* **215**, 355–359.
29. Lanks, K. W. & Kitchin, F. D. (1972) *Nature* **226**, 753–754.
30. Clarke, B. (1970) *Nature* **228**, 159–160.
31. Sneath, P. H. A. (1966) *Theoret. Biol.* **12**, 157–193.
32. Holmquist, R., Jukes, T. H. & Pangburn, S. (1973) *J. Mol. Biol.* **78**, 91–116.
33. Fitch, W. M. & Markowitz, E. (1970) *Biochem. Genet.* **4**, 579–593.
34. Fitch, W. M. (1972) in *Haematologie und Bluttransfusion* ed. Martin, H. (J. F. Lehmanns Verlag, Munich, Germany), pp. 199–215.
35. Li, S. L., Denney, R. M. & Yanofsky, C. (1973) *Proc. Nat. Acad. Sci. USA* **70**, 1112–1116.
36. Ohno, S. (1970) *Evolution by Gene Duplication* (Springer-Verlag, Berlin).
37. Ohno, S. (1973) *Nature* **244**, 259–262.
38. Nei, M. (1969) *Nature* **221**, 40–42.
39. Ohta, T. & Kimura, M. (1971) *Nature* **233**, 118–119.
40. Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. C., Mercola, D. A. & Vijayan, M. (1971) *Nature* **231**, 506–511.
41. Ohta, T. (1972) *J. Mol. Evolut.* **1**, 305–314.
42. Kihara, H. (1947) *Ancestors of Common Wheat* (in Japanese) (Sōgensha, Tokyo).