

Natural Selection and the Origin and Maintenance of Standard Genetic Marker Systems

FRANK B. LIVINGSTONE

Department of Anthropology, The University of Michigan, Ann Arbor, Michigan 48109

KEY WORDS: Natural selection, Polymorphism, Human genetic markers, Human genetic variation, Evolutionary theory, Neutral mutation

ABSTRACT Natural selection has always been assumed to be the major force of evolution, but its presence has been difficult to demonstrate. A review of the evidence for selective differences among genotypes for most human genetic polymorphisms indicates there is little of a direct nature. Indirect theoretical evidence, however, seems to support a major role for natural selection, and it does not seem to support the hypothesis that most amino acid substitutions within the human species are neutral. Among small isolates, most of the gene frequency differences are most likely due to genetic drift or the founder effect, and the principal counterbalancing force is gene flow or migration. But genetic differences among the major human subdivisions do not seem to be due to the same interacting forces. One reason for the inability to detect selection has been an oversimplified view of its operation, which assigns genotypes a constant fitness in every generation. Many recent theoretical developments of more complicated kinds of selection may lead to a resolution of the problem and suggest better interpretations of the enormous amount of data on human genetic variation that is rapidly accumulating.

The extraordinary advances in biochemical genetics in the last two decades have uncovered an astonishing amount of genetic variation in human populations with respect both to the great number of alleles present at some loci and to the high frequency of polymorphic loci. A locus is now generally defined as polymorphic if at least two alleles are present in frequencies greater than 1%. Current studies based on blood group and enzyme loci estimate that about 28% of human loci are polymorphic (Hedrick and Murray, 1978). Some genes, such as the four major HLA loci, have an extreme amount of polymorphism, with an average of over ten such alleles (Bodmer and Bodmer, 1978), whereas other genes, such as the two major hemoglobin loci, have over 300 known genetic variants but with very few occurring in polymorphic frequencies (McKusick, 1978; Weatherall and Clegg, 1979).

Given the amount of morphological variation in the human species, perhaps this amount of genetic variation should not be surprising; but it has focused attention on the possible explanations for polymorphism. All such explanations have been based on population genetics theory, but with varying emphasis on the various evolutionary forces that can account for genetic similarities and differences among populations. The issue has been phrased as the "neutralists vs. selectionists," which points out that it is primarily the role of natural selection that is at the center of the controversy.

Despite the wide acceptance of Darwinian theory and its major concept of natural selection as the principal determinant of evolutionary change, it has been difficult to obtain direct evidence for natural selection in operation, and there are few cases where it has been measured. Most of us can agree with Lewontin (1974:199) that "anyone who has taught genetics for a number of years is tired of sickle cell anemia and embarrassed by the fact that it is the only authenticated case of overdominance available." We are also likely to be "tired" of industrial melanism in moths as the other overworked case of demonstrable selection that led to a marked evolutionary change in a species. Even so, Darwinian evolution is still assumed to be the major process that has led to genetic diversity among species.

The concern with natural selection is justified, because it would seem to be necessary to determine the effect of this cause of genetic change before other explanations can be explored. Although this theoretical exercise is frequently put in an either/or context, it more realistically depends on the relative magnitude of the forces of evolution and on the units of analysis. Thus, natural selection may well be the major determinant of the gene frequency differences that occur among the principal divisions of the human species, but other forces can have considerable influence on the genetic variation among individual isolates.

Recently there has been a change in the conceptualization of the forces of genetic change, with a growing consensus that these are mutation, selection, and population structure. Previously gene drift and gene flow were considered as separate entities, but drift is a function of isolate size (N), and gene flow or migration is a consequence of mating patterns and population dispersion. The amount of inbreeding is also a function of both the mating system and isolate size, so that all these other forces can be included in the general term, population structure. This conceptual change is due primarily to a shift in research away from the study of a single isolate to the study of several closely related isolates or populations. In addition, all the parameters of population structure affect all loci equally, so that great differences in variation among loci within a given set of isolates are probably due to the other forces, and especially to natural selection.

Many years ago, Wright (1951) developed the theoretical basis of steady state gene frequency distributions by which the relative effects of the various parameters of these distributions could be evaluated. Mutation, selection, and gene flow are all systemic pressures that tend to equate the gene frequencies of a set of isolates. If their total effect is less than $1/2N$, then the distribution will be U-shaped, and there will be considerable fixation and extinction of alleles. On the other hand, if their total effect is much greater than $1/2N$, then the distribution will be very peaked, and the mean equilibrium value will be determined by the systemic pressures. Of course, the variation around the equilibrium will be due to drift, but it will be very small with large systemic pressure.

Today most human isolates are greater than 1000 in size, with some obviously much larger. In the past, human isolates were undoubtedly much smaller on the average, and perhaps averaged as low as 50 or 100 among hunters. According to steady state theory, today selection and other systemic pressures would have to be greater than .0005 to have a major effect, whereas in the past pressures of .01 (or twenty times today's maximum value) would have been significant. In both cases these values seem very small and less than most estimates of gene flow or migration among human isolates. The value of $1/2N$ is debatable, but even if the threshold were increased to $1/N$ or even $2/N$, which would make the steady state distribution much more peaked, the amount of selection or other systemic pressures necessary to be effective is still quite small. Furthermore, selection of this magnitude is practically unmeasurable in most human populations for at least two reasons. First, selection is literally "swamped" by gene flow, which is at least .05 for most human isolates, and second, the size of the sample necessary to detect fitness differences that could have an important effect on gene frequencies is often larger than the isolates being studied, or frequently too large to be collected.

Given these difficulties, it is easy to see why the acceptance or recognition of selection as a major force determining genetic variation within the human species has been achieved reluctantly. Until the recognition of the importance of the problem of maintaining polymorphisms in the 1950s, the classical view of genetic variation assumed that most loci were monomorphic and that any deviation from the normal allele was due to deleterious mutations. According to this view, the process of evolution was the replacement of one allele by another, so that polymorphisms were considered either to be a locus in the process of replacement of alleles, or, more significantly, to be nonadaptive or neutral substitutions. Lewontin (1974) has characterized as the neoclassical view the recent comparable position that most of the amino acid substitutions that occur between species are neutral.

The blood groups were the major genetic polymorphisms known at that time for the human species, and despite the known operation of selection through maternal-fetal incompatibility, they were generally considered to be nonadaptive. There was some speculation that perhaps the B blood group allele was a later mutation that was in the process of spreading throughout the species. One of the most influential books in anthropology in this era was W.C. Boyd's (1950) *Genetics and the Races of Man*, which expressed the nonadaptive view in several ways:

Among the racial characteristics which we would be tempted to pick out at the present time as non-adaptive, there are certain serological features of the blood, such as the genes O, A, B, M, N, etc. (p. 27);

or

The blood grouping characteristics, with the possible exception of the Rh factors, seem to have been exonerated from being of much selective value in so far as we have been able to test them (p. 150);

and finally,

When we come to the second agency which might have modified blood group gene frequencies in the world, namely, natural selection, we shall be obliged to return, on the whole, a verdict of not proved. It is very difficult to establish that a character has no selective advantage, whatever, and, as Fisher has shown mathematically (10), selective advantages which seem at first sight very small may nevertheless be quite sufficient to bring about evolutionary modification of the characteristics of a species. We may tentatively suggest that *if* selection does act on the blood group characteristics, there is some reason to think that B is somewhat favored over O and A, since the present distribution of blood group frequencies suggests that the frequency of the B characteristics has increased markedly in Asia and Europe within fairly recent times geologically, or even within historical times (3). However, this increase could probably be accounted for by other factors, and is no more than a slight indication that some advantage might inhere in the possession of the blood group B (p. 335).

These quotations show that the possibility of natural selection being an important factor was recognized despite the small values it might have, but even so there is a reluctance to ascribe any meaningful role to selection in determining genetic variation. They also show that selection was usually considered in terms of replacement of one allele by another, with no consideration of how selection could maintain genetic variability. Boyd's book has several appendices that give the equations for gene frequency change and genetic equilibria, but those for a balanced polymorphism with intermediate frequencies of two alleles are conspicuously absent, although they had been known for twenty years.

Biologists other than anthropologists had the same reluctance to accept natural selection as an important determinant of human genetic variation and seemed to have the same desire to find other explanations. In a long essay on human variation, Dobzhansky (1950) discussed polymorphism in terms of a balance of mutation rates but with no mention of selection. He concluded: "The blood-group genes are, it appears, a good example of such 'neutral' genes. Here is a virgin field for anthropological

studies" (p. 133). As the alternative explanation to selection, Dobzhansky concluded, "Thus the mutation pressure is enough to account for the existence of several blood groups in human populations" (p. 128), and the variation in gene frequencies was explained as: "In the case of blood groups genetic drift appears to be the probable agent of race differentiation" (p. 136).

Mutation pressure is no longer considered a reasonable explanation since back mutation is rare. Most recent models assume each mutation is unique, and this is a closer approximation to the biochemical realities of DNA. Nevertheless, the same disregard of natural selection is apparent in most recent attempts at phylogenetic reconstruction by gene frequency similarities and differences of species and racial relationships. And, although most studies do not explicitly say so, the underlying forces are assumed to be neutral mutation and genetic drift, for which Kimura (1968) and King and Jukes (1969) provided the first models.

Two developments in the 1950s led to the recognition of the importance of natural selection in determining genetic variability. One was the rediscovery of the association of the ABO blood groups with various diseases, all of which have recently been compiled by Mourant et al. (1978). The other was the discovery that two deleterious mutants, sickle cell anemia and thalassemia, were very widespread and attained very high frequencies in many human populations. The first two diseases to be associated with the ABO blood groups were peptic ulcer and stomach cancer, and it was quickly realized that these diseases have little evolutionary significance since they affect fitness only slightly by occurring primarily after reproduction is completed. However, they did bring the realization that the blood groups were not nonadaptive, and these findings led to other investigations on diseases such as viruses and other infectious agents that did have great effects on fitness (Reed, 1975; Vogel, 1975). On the other hand, the enormous compendium by Mourant et al. (1978) results in few concrete conclusions as to how infectious and other kinds of disease have influenced the gene frequencies of the blood groups.

Haldane (1949) is generally credited with suggesting that malaria might be the balancing factor for thalassemia, although this is not mentioned specifically in the text of his paper, but by Montalenti in a comment on the paper. Haldane only suggested that infectious disease was a likely balancing force for the ABO blood groups! Allison (1954) provided the first concrete evidence for falciparum malaria being responsible for the high frequencies of the sickle cell gene. As for any major advance in science, others had previously suggested a possible relationship between the sickle cell trait and malaria (Beet 1946, 1947; Raper, 1949; Brain, 1952a, b), and there was some evidence of a geographical association between thalassemia and malaria (Vezzoso, 1946; Bianco, 1949). But with Allison's evidence and subsequent discoveries of the much lower mortality from cerebral malaria of sickle cell trait carriers, balanced polymorphism was finally demonstrated for a human locus (Motulsky, 1975).

It is easy to see why the sickle cell polymorphism was the first to be discovered and why it may be the only one to have conclusive evidence of a direct nature ever found. The homozygote for the sickle cell gene has a fitness that is very close to 0 in most of Africa. With the highest known frequencies of about .2 in Africa, the normal homozygote must have a fitness of about .75 relative to the heterozygote if these frequencies are close to equilibrium. Thus, the fitness differences between genotypes are almost as great as they could be, and these differences are obviously much easier to measure than those that are presumably more common in evolution. If, as was pointed out previously, fitness differences of .01 or less would be important determinants of genetic variation, then such differences would occur and would be unmeasurable by the methods and sample sizes of most investigations on human populations. That the sickle cell polymorphism is the only one with good evidence in its favor should thus not be surprising. Reed (1975) has reviewed the studies that attempted to measure the "maximum amount of undetected selection" that could exist for various loci, and has concluded that selective effects as large as 10-20% could go undetected.

The sickle cell-malaria association rapidly became one of the dominant paradigms of evolutionary genetics, and its many implications have been explored in recent years with varying degrees of success. First, this association implied that malaria must have

been an extremely important selective factor in the evolution of many human populations in the tropical and subtropical regions of the Old World, and hence there may be many other polymorphisms adapted to it. Second, it pointed to the important role of infectious disease in general as a selective force. Third, many deleterious genes that are found in restricted numbers of populations in frequencies higher than could be maintained by mutation may well be balanced polymorphisms and perhaps balanced by infectious disease. Finally, it raised the strong possibility that many of the known genetic polymorphisms such as the blood groups and the great many serum proteins and red cell enzymes that have been discovered in the last twenty years were balanced polymorphisms. These implications have all been explored in recent years, and these investigations have led to an increased understanding of human genetic variation. These implications will now be considered in turn.

At the same time that the relationship between sickling and malaria was being investigated, other abnormal hemoglobins were being discovered at an accelerating rate. The different forms of thalassemia were also being elucidated and were found to vary among human populations. By obvious analogy with the sickle cell gene, these abnormal variants that were found in high frequencies were considered to be due to a resistance to malaria on the part of the heterozygotes. Although there are now over 300 known variants of adult hemoglobin, very few attain polymorphic frequencies in a large number of human populations. In addition to hemoglobin S and β -thalassemia, only hemoglobin C in West Africa and hemoglobin E in Southeast Asia are widely distributed in very high polymorphic frequencies. However, many other hemoglobin variants are found in much lower but still polymorphic frequencies in much more restricted populations, for example, hemoglobin D_{Punjab} in India, hemoglobin G_{Accra} in Jamaica, and hemoglobin O_{Indonesia} in Sulawesi (references in Livingstone, 1967, 1973).

For years it has been known that thalassemia was a heterogeneous syndrome due to many different mutations. The two principal kinds were α and β thalassemia, which acted as alleles of the α and β structural variants. With recent techniques using restriction endonucleases, the specific alteration in the DNA structure can be detected (Dozy et al., 1979), so that the variations of thalassemia can be defined and are now known to be many more than expected. Dozy et al. (1979) have found that about 27% of American blacks are carriers of a mild type of α thalassemia, which seems to be another polymorphism maintained by malaria.

Hemoglobins C and E do not have the severe effects of hemoglobin S or thalassemia when homozygous, which has led to suggestions that these are not balanced polymorphisms. Cavalli-Sforza and Bodmer (1971) developed a computer model of the diffusion of the hemoglobin C allele in West Africa that has it replacing both the hemoglobin S and hemoglobin A alleles in between 200 and 500 generations. Das et al. (1975) have reported a frequency of over 50% for the hemoglobin E allele in some populations in Assam, and postulate that it may replace the normal allele under conditions of severe malaria. In both cases, however, it seems that the heterozygote still has the highest fitness, so that a polymorphism would still exist; although with the CC or EE homozygote having a higher fitness than the normal homozygote, the equilibrium frequency for the abnormal variant would be greater than .5.

Despite the fact that the geographical distributions of these other hemoglobin variants strongly suggest that their high frequencies are due to a resistance of their carriers to malaria, there is little direct evidence for this difference in fitness (reviewed in Livingstone, 1971, and Motulsky, 1975). The recent development of techniques to grow malaria parasites *in vitro* has provided very good evidence for the hemoglobin S resistance to falciparum malaria (Friedmann, 1978; Pasvol et al., 1978; Roth et al., 1978) and for some protection by increased levels of hemoglobin F (Pasvol et al., 1977), so that in the future, evidence for the other hemoglobins being resistant to malaria may be forthcoming.

The other polymorphism that has been associated with malaria is the glucose-6-phosphate dehydrogenase deficiency. Over 100 variants of the G6PD enzyme are now known, but with the exception of Gd A+ throughout tropical Africa, most of the variants in polymorphic frequencies are deficient in enzyme activity. However, there is recent evidence that there may be much more variation in this enzyme than previously supposed (Modiano et al., 1979). There is some evidence that the female

heterozygote may have a selective advantage with regard to malaria (Bienzle et al., 1972), but this has been the subject of a recent controversy (Martin et al., 1979; Luzzatto and Bienzle, 1979). Nevertheless, there is evidence that G6PD deficient red cells are oxidant-sensitive which suppresses malaria infection (Eaton et al., 1976), and Friedmann (1979) has recently shown that G6PD deficient red cells and those of both α and β thalassemia are not as good a medium for falciparum malaria parasite growth as normal cells.

In addition to G6PD, many other red cell enzymes are known to have genetic variants. Some of these, such as pyruvate kinase, have very different frequencies in blacks (Chern and Beutler, 1976), and there is some evidence of an association of pyruvate kinase activity with malaria (Martin et al., 1978). Finally, other red cell abnormalities such as ovalocytosis are found in high polymorphic frequencies in some populations with endemic malaria, which seems to be the most likely selective agent increasing their frequencies (Serjeantson et al., 1977).

Investigations have been done on the relationship of other genetic polymorphisms such as haptoglobin and malaria, and there have been many studies on the blood groups and malaria, but most of this work is inconclusive (Livingstone, 1971). More recently, however, the Duffy blood group system has been shown to be the site on the red cell membrane of the invasion of the cell by *Plasmodium vivax* (Miller et al., 1976). Individuals who are negative for both Fy^a and Fy^b are in the great majority in African populations and have been shown to be almost completely immune to vivax malaria infection (Miller et al., 1978; Spencer et al., 1978; Welch et al., 1977).

There has thus been considerable success in implicating malaria as a selective factor determining the frequencies of many genetic polymorphisms. What seems puzzling, however, is that no other infectious disease has been shown to be the principal factor determining the frequency of a genetic polymorphism. Perhaps this is due to the higher and more consistent mortality rate of malaria or perhaps to its intimate relationship to the red cell, which has another major function; but there is no other similar association. There are, however, many associations between infectious disease and the blood groups (Vogel, 1975; Mourant et al., 1978). Two such diseases, smallpox and influenza, have certainly had selective effects, but their role in determining the variation in gene frequencies is problematical. Similarly, there is a rapidly accumulating mass of evidence for the strong association of various HLA antigens with various diseases (summarized in Bodmer and Bodmer, 1978), but their significance for genetic variation is unknown, and the antigen is often more common in populations where the disease is also more frequent. This absence of any one disease-one polymorphism association for either antigen system seems expectable since the antigen systems are part of the antigen-antibody immune system that protects the individual against all diseases and other foreign tissue invasions.

Although the direct evidence for the selective effect of infectious diseases is slim, if almost nonexistent, many genetic polymorphisms have been thought to be balanced by disease selection. These polymorphisms are the few lethal or very deleterious alleles that are found in some, usually restricted human populations. One of the first to be suggested to be balanced by infectious disease was schizophrenia (Huxley et al., 1964), although the genetic basis of this disease is still not known. But Tay-Sachs disease, which attains a gene frequency of about .02 among Eastern European Jewish populations, and cystic fibrosis, which occurs in frequencies of up to .02 in many large European populations and in a frequency of .05 in one region of France (Bois et al., 1978), have been generally accepted as balanced polymorphisms, with infectious disease as the presumed selective factor (Knudsen, 1979). Other polymorphisms such as the protease inhibitor allele, Pi Z, the homozygosity for which is strongly associated with liver and lung disease, and phenylketonuria are also found in frequencies of .02 to .04 in some European populations.

There are other instances of polymorphic frequencies of deleterious alleles in more restricted and smaller human populations that in many cases are at the same locus as those mentioned above. For example, cystic fibrosis is found in a frequency of .08 in Southwest Afrikaners (Super, 1975) and Tay-Sachs disease occurs in a religious isolate in Berks County, Pennsylvania, at a frequency close to .02 (Kelly et al., 1975). Because

of their small size and their origin from a relatively few founders, these latter populations could be expected to have high frequencies of deleterious alleles due to the founder effect, but for large national populations this explanation seems much less probable. Nevertheless, the Eastern European Jewish populations were founded by very few individuals who escaped to Lithuania from the persecutions of the Roman Empire. Since they apparently have the greatest number of deleterious genes in high if not polymorphic frequencies—23 is the current estimate (Chakravarti and Chakraborty, 1978)—the founder effect seems to be a more likely explanation for their increased genetic load than for other large populations (Fraikor, 1977; Livingstone, 1969, in press). However, the opposite view, that these frequencies are not due to the founder effect but to heterozygote selection, seems to be the majority view (Chakravarti and Chakraborty, 1978; Wagener et al., 1978).

Infectious disease has been the principal factor advanced as balancing these polymorphisms. The ghetto environment of Eastern Europe was implicated as causing a high incidence of tuberculosis, typhoid fever, typhus, and others, and the hexosaminidase deficiency, which is the genetic defect in Tay-Sachs disease, has been proposed as providing some resistance to disease. In replying to Chase and McKusick's (1972) suggestion that Tay-Sachs disease was due to the founder effect, Myriantopoulos et al. (1972), cited evidence that the heterozygotes may be resistant to tuberculosis. But this evidence was a lower death rate from tuberculosis in the areas of Eastern Europe where the Tay-Sachs gene frequencies are highest (Myriantopoulos, 1972). But if selection by tuberculosis is responsible for the Tay-Sachs polymorphism, it should be more endemic where Tay-Sachs is highest.

Because of its widespread polymorphic frequencies in Northwest Europe, cystic fibrosis is also assumed to have been maintained by heterozygote advantage to some disease. The sera of individuals with cystic fibrosis agglutinate the *Proteus* OX19 bacteria, so that selection by typhus has been suggested (Stuart and Burdon, 1974). But typhus has been more endemic in Eastern Europe and the Mediterranean than in Northwest Europe. There has also been a reported association of cystic fibrosis with tuberculosis (Crawford, 1975), but no conclusive evidence has been forthcoming, and these hypotheses only have plausibility because of the presumed problem of these polymorphic frequencies.

The major geographic subdivisions of the human species vary considerably in the deleterious mutants that approach polymorphic frequencies in their populations. Phenylketonuria is common in Europe. In Japan acatalasia attains a frequency of .017 in some populations but is very rare elsewhere (Neel et al., 1963). At equilibrium one would expect all large populations to have similar frequencies of common deleterious mutants. This raises the possibility that the enormous expansion of some human populations over the last 1000–2000 years is responsible for these genetic differences. After rapid expansion from a limited number of founders, a lethal or deleterious allele that happened to attain high frequencies due to the founder effect will persist for a very long time due to the slow selection against them in large panmictic populations. I have simulated population growth with various estimates of the parameters of founder size, selection, and growth rate, and the results indicate that many of these polymorphic frequencies of deleterious alleles could be due to the founder effect (Livingstone, in press). For example, beginning with a population of 50 individuals who would be heterozygous for at least 50 lethal equivalents, a doubling of the population for 6 generations would result in approximately 50% of these alleles attaining polymorphic frequencies, after which they will decrease slowly. The extremely widespread distribution of such genetic traits as cystic fibrosis or the protease inhibitor allele, Pi Z, in Northwest Europe seems to mitigate against such an explanation, but it is noteworthy that these populations have probably undergone the most rapid expansion of any human populations in the last 1000 years.

For many deleterious alleles that occur in polymorphic frequencies of up to .04 in large populations, it thus seems that the verdict of selection vs. founder effect is still not in. No concrete evidence has been published for any association with disease, and the demographic history of the last 1000 years leaves open the possibility of the founder effect. The high frequencies of Tay-Sachs disease and Niemann-Pick disease

in French Canadians could be examples of the same process at a later time in the New World (Andermann et al., 1975; Winsor and Welch, 1978). Of course, selection to maintain a frequency of a lethal of .03 would only be about .03, which is essentially unmeasurable even if it still exists. Nevertheless, selection seems to be most likely, according to most geneticists, at least for widespread alleles like cystic fibrosis (Knudsen, 1979).

The last implication of the rediscovery of selection with the sickle cell-malaria association is that many of the large number of recently discovered polymorphisms are very likely due to selection, and probably to selection for heterozygotes. Even polymorphisms that have been known for a long time, such as the blood groups or the PTC taster, are now generally assumed to be balanced by heterozygote advantage. Thus, Cavalli-Sforza and Bodmer (1971:213) state that "if the ABO polymorphism is stable, some selection effectively favoring heterozygotes must counterbalance incompatibility selection."

Various types of electrophoresis led to the discovery of over 300 variants of human hemoglobin, and with further refinements have resulted in the discovery of the enormous number of polymorphisms in red cell enzymes, serum proteins, and in other body tissues and fluids. In addition, advances in immunological techniques have led to the discovery of the large amount of genetic variation in the immunoglobulins and in the histocompatibility antigens. Giblett (1969, 1977) has outlined the genetic characteristics of these polymorphisms, and Mourant et al. (1976a) have tabulated their frequencies in human populations. As the title of Giblett's (1969) book and of this review indicate, all these genetic polymorphisms have become known as markers, and the question arises as to what they mark. Giblett (1977) has outlined the many clinical and genetic uses of markers, such as chromosome mapping, genetic mosaicism, and the problem of tumor origins, but for the fields of population genetics and anthropology, there has been little discussion about the primary use of markers. That the possession of a similar marker or similar frequencies of a particular marker is generally used to indicate a genetic relationship between two populations is so obvious as to require little comment in the literature. For example, Lucarelli et al. (1979) state that the placental alkaline phosphatase polymorphism has great anthropological value as a genetic marker because it differentiates the three major ethnic groups of mankind and because it supports the hypothesis that the Sardinians can be considered as a differentiated branch of Caucasians, whatever that means.

This use of the term, marker, is applied by most geneticists and anthropologists to all polymorphisms that vary among populations, and Crawford (1973) lists "nearly 50" which he defines as "discrete, segregating, genetic traits which can be used to characterize populations by virtue of their presence, absence, or high frequency in some populations and low frequency in others" (p. 38). Although this definition is innocuous enough by itself, it leaves in abeyance the question of why two populations have similar or different gene frequencies. Obviously this question should be considered within the framework of population genetics theory, but it is simply ignored in most investigations. The number of studies reporting frequencies of genetic polymorphisms in various human populations has been increasing and is reaching extraordinary proportions. But most have only reported the frequencies, with perhaps a suggestion as to which alleles are distinctive for that population or are markers. Mourant et al. (1976a), compiled all the known data for most genetic polymorphisms in man, and the tables fill almost 800 pages, whereas the discussion or attempted explanation is only 60 pages, and the attempt at synthesis a mere six. I do not mean to disparage Mourant's outstanding effort in bringing together all this work, but only to show the enormous size of the data bank which has no adequate explanation.

The computer has made the analysis of large masses of data an easy exercise. All of physical anthropology has succumbed to the allure of the computer, so that principal components analysis, cluster analysis, and other sophisticated statistical techniques have been applied to everything from ankle bones to fingerprints to show who is more related to whom. Cavalli-Sforza and Edwards (1963) first developed an analytic method using gene frequency differences that was supposed to reveal phylogenetic relationships, which they expressed as an evolutionary tree. Since then their model or similar

ones have been applied to most of the data on genetic polymorphisms. In most cases the frequencies of the hemoglobin variants and the G6PD deficiency are not used, since these loci are known to be related to malaria and due to selection. This maneuver obviously assumes that the other polymorphisms are not influenced by selection.

Cavalli-Sforza and Edwards (1963) actually published two evolutionary trees, one based on genetic polymorphisms and the other on anthropometric traits. These differ considerably, with the genetic one showing closer ancestries for the Africans and Europeans as opposed to the North and South Asians and Pacific peoples, whereas the anthropometric tree shows the Africans to be most close to the Southern Asians, and the Europeans and the North Asians to be classified together. The latter obviously shows selection for skin color, nose width, curly hair or other anthropometric traits that are probably selected for in a tropical climate. Although there is no similar selection known for the distribution of the genetic polymorphisms, there seems to be no apparent reason to disregard such an association or to assume it is due to common ancestry.

Nevertheless, the major use of genetic markers has been to construct evolutionary trees, but these constructions have added little to our knowledge of the course of human evolution in the sense of detecting phylogenetic relationships that were not known from other data. Fitch and Neel (1969) used the blood group and other polymorphic frequencies known at that time for South American Indians to construct an evolutionary tree for the entire continent. Some close relationships seem plausible and others are known on other data, but the Guaymi tribe of Panama clustered closest to the Yanomama of Southern Venezuela and Brazil. Further investigation with more loci has resulted in a different tree, with no close relationship between these two tribes (Spielman et al., 1979). This shows that the actual polymorphisms used can have a considerable effect on the resulting tree, and this in turn implies that the individual loci are not measuring the same thing.

The HLA loci have been the subject of a great number of studies that show many associations with disease (summarized in Bodmer and Bodmer, 1978). However, the workshops that have been devoted to their variation among human populations have been summarized by tree diagrams purporting to show phylogenetic relationships. Piazza and Viganotti (1973) used both cluster analysis and principal components analysis with similar results. For all polymorphic loci, Northern Italians are closely associated with Lebanese and Turks; Basques and Icelanders are very close and then closest to Yemenites; while many Sardinians are closely related, others are more related to Koreans and Lapps. Tibetans and Australian Aborigines are very closely related and then more closely related to the Warao, a small tribe in South America, while the Eskimos are more closely related to some New Guineans than are other New Guineans. For the HLA loci alone the evolutionary tree shows that the Basques are most closely related to the Congolese and then closer to the Icelanders than to other Europeans, while the Lapps are most close to the Warao, who in turn are closer to European and some African populations than to other New World populations. These "phylogenetic" relationships seem highly unlikely, to say the least. And I think it fair to say that if an archaeologist had made similar suggestions on the basis of some similarity in an esoteric artifact, or an ethnologist on some similarity in mythology, they would be laughed out of court. That such absurd "phylogenies" have been published with regularity seems to be a futile use of the enormous accumulation of data on genetic polymorphisms. As Morton (1972:132) has remarked, "phylogenetic trees are like flower arrangements: it is enough that they are pretty, without asking that they be meaningful."

On the other hand, studies of recent population history such as those by Crawford et al. (1976) on the Tlaxcaltecs or the many mixed groups in the United States can obviously show close genetic relationships that are reflected to a comparable degree in other morphological traits or in the known history of the population. Although Harpending (1974) considers this use of genetic data for historical reconstruction to be its most promising future, I cannot see how it adds anything to either genetic understanding or to regional history. Usually the regional history is known in detail, and to show that the genetic data accord with it seems of little import, since these

concordances only exist because of the short time span involved. But with a longer time span, which gives selection sufficient time to become effective, discordance among loci will develop, since selection acts on individual loci whose functions and relationships to the environment can be very different.

This is evident in the original evolutionary trees for anthropometric and genetic traits constructed for the human species by Cavalli-Sforza and Edwards (1963). Recently, however, Szathmary and Ossenberg (1978) have compared genetic and craniometric traits among Eskimos, Indians, and some Asian populations and shown that there is concordance which is statistically significant and that the Eskimos are more closely related to some Indians than Asians. Nevertheless, there are many obvious discordances and relationships that are highly dubious. For example, for genetic traits the Apache are very distant from the Navajo and more closely related to the Asiatic Mongoloids, while they are closest to the Navajo for cranial traits. There are many other phylogenetic conclusions that could be drawn from their analysis, so that Szathmary and Ossenberg (1978) have selected those that seem plausible or noteworthy. They could have also concluded that the East Greenland Eskimos are closest to the North Alaskan Eskimos, and the West Greenland Eskimos are closest to the South Alaskan Eskimos.

After the rediscovery of the importance of selection, Boyd (1963) recognized the same races of man that he had previously, but he stated that selection was the major factor responsible for their differentiation. However, Edwards and Cavalli-Sforza (1964:74) disagreed and stated that "Boyd's comment, 'unless the blood groups are adaptive, they are not going to be very useful in racial classification,' is not acceptable." But the problem of why a marker gene is present in a large segment of the human species is still unresolved. Most recent models that use a phylogenetic analysis of gene frequencies, such as Nei and Roychoudhury's (1974), implicitly assume a rapidly expanding, branching population history of *Homo sapiens* that began with a few founders 50,000 or more years ago. This model contrasts with the view of Coon (1962) and others that the geographical variants of man have been in their relative positions in the Old World for most of the Pleistocene, and the phenotypic variation is due primarily to natural selection. Undoubtedly both rapid expansion and populational equilibrium have characterized man's history, but the widespread distribution of a genetic marker over thousands of miles and hundreds of populations could only be maintained by natural selection if there were relative stability of the populations. For examples, the Diego blood allele, Di^a , is found in polymorphic frequencies from India through Asia, and in the Amerindians, and the albumin variant, Naskapi, is found in Amerindians from New Mexico to Labrador in low but polymorphic frequencies. There have been population movements within these areas of their distributions, but there has been considerable stability; and it would be impossible for an allele to spread through these populations without having a selective advantage.

Although an enormous amount of the analysis of recent polymorphisms has been of the kind already discussed (that disregards selection and thus does not follow the sickle cell-malaria paradigm), there have been many investigations to detect selection acting on these polymorphisms. The most obvious are the direct attempts to show differential liability to infectious or other diseases on the part of the different genotypes, but many indirect methods have been devised to detect selection. Mourant et al. (1978) have compiled most of the data on association between genetic polymorphisms and disease, and the investigations continue, with associations between polymorphisms and disease being tested for goiter (Harrison et al., 1976), psoriasis (Beckman et al., 1977), allergic diseases (Brachtel et al., 1979), diabetes (Lucarelli et al., 1978), and leprosy (Hitzeroth et al., 1978). There are also many highly significant associations between HLA haplotypes and disease (Bodmer and Bodmer, 1978). However, all this work on disease associations is very inconclusive with regard both to the significance of the associations (since in many cases some studies find one but others do not) and to the effect of the various diseases on the frequencies of the alleles with which they have been found to have significant associations. While for the sickle cell-malaria association the differences in fitness are determined primarily by the one disease, falciparum malaria, for other genetic polymorphisms the association is far

more complex. Hence the role of any single disease in determining the gene frequencies is uncertain. In addition, the type of selection involved does not seem to be simple heterozygote advantage, as in the sickle cell-malaria case.

Many other indirect kinds of evidence have been used to detect, if not measure, selection. One is the correlation of gene frequencies with various environmental or geographical variables, and the others mostly attempt to estimate the various parameters of gene frequency change from the distributions of gene frequencies in human populations. Hiernaux and Froment (1976) correlated all the known frequencies for genetic polymorphisms in African populations with many climatic variables and found few significant correlations, and none that seemed to indicate selection conclusively, with the exception of the hemoglobins and G6PD deficiency. There have been, however, studies of individual polymorphisms based on all human data that show many significant worldwide correlations with climate. These include the transferrins (Walter and Bajatzadeh, 1971), acid phosphatase (Ananthakrishnan and Walter, 1972), human cerumen types (McCullough and Giles, 1970), and the Gc serum protein (Mourant et al., 1976b). The Gc protein is known to be the plasma carrier of vitamin D (Daigher et al., 1975), and hence the allelic differences appear to be associated with the amount of sunshine.

Although these correlations are good evidence for selection, there seems to be no method to estimate directly its magnitude or effect. But if the other parameters of gene frequency change can be estimated or assumed, then with genetic models it is possible to estimate how much selection is required to obtain the clines and gene frequency distributions within the human species. However, most models of gene frequency distributions have been developed to show that selection is not a major determinant of gene frequency differences among human populations. Beginning with Kimura's (1968) suggestion that most amino acid substitutions are neutral, there has been a plethora of models to demonstrate neutrality. Ewens and Feldman (1976) have reviewed most of these attempts and conclude that "all models fail to demonstrate neutrality." Their conclusion was based on the many mathematical assumptions these models have to make that could be shown to be false or at least very unlikely. One of these models was developed by Yamazaki and Maruyama (1972) and showed that the distribution of heterozygosity varied with various kinds of selection or with no selection. The distribution for over 400 polymorphisms seemed to accord with the neutral model. Later, Yamazaki and Maruyama (1974) expanded their sample to 1045 polymorphisms, and it still conformed to the neutral model. However, the distribution for 26 human blood group loci did not, but rather seemed to indicate "balanced selection." Although Yamazaki and Maruyama claim that the distribution of heterozygosity is independent of population structure, Neel (1978) has shown that the distribution varies significantly between tribal populations and larger national ones, and Beals and Kelso (1975) showed that blood group heterozygosity is associated with cultural level and increases in large modernized societies.

The distribution of gene frequencies in human populations thus does not accord with neutralist theory, and this is particularly true for some loci that have the greatest amount of known data on gene variation, particularly the blood groups and the hemoglobin loci. Given the great number of hemoglobin variants that have been discovered in human populations, it is contrary to theory that no variant that is generally accepted as an example of the process of "neutral" amino acid substitution has been found in high frequencies. In fact, all the polymorphic variants have some selection against them and are the best examples of balanced selection. This situation caused Haigh and Maynard-Smith (1972) to develop a model of the bottleneck effect and propose that the human species went through a bottleneck of several thousand years back in the Pleistocene. Boyer et al. (1978) have shown that the mutations for the β hemoglobin locus occur at random along the entire polypeptide, and that 95% of these mutations are "unacceptable" since they are very deleterious when homozygous. The other 5% of "benign" mutations do occur but have not attained high frequencies in any populations. This is also evidence against neutral evolution. However, Boyer et al. (1978) point out that the differences in hemoglobin structure among mammalian species are these "benign" mutants.

The same contradiction seems to occur for many other proteins. For example, Sarich (1977) has assumed that the evolution of the genetic structure of albumin occurs at a constant rate and can be used as an evolutionary "clock." Although his own work is based on immunological differences, the model is similar to that proposed by Kimura (1968) for the constant change in amino acids. But there seems to be little evidence in the distribution of albumin variants in human populations for this constant process of neutral replacements. Albumin is not polymorphic in most human populations, and the few variants that are found in polymorphic frequencies among Amerindians and some other populations (Schell et al., 1978) occur at very low frequencies. Thus, no neutral replacement seems to be occurring, and since the polymorphisms that are found are widely dispersed among relatively unrelated populations, their distributions seem to require some selection to maintain them at these low frequencies.

The amount of heterozygosity or the gene frequencies of all polymorphisms when considered together do not provide much evidence for selection, but neither do they provide good evidence for the absence of selection. The comparison of individual polymorphisms appears to offer the most hope of detecting selection, although to date the methods are still the object of much debate. These methods are based on the fact that the parameters of population structure should affect all loci equally. If estimates of these parameters, most especially of gene flow and gene drift, derived from different polymorphic loci vary significantly, then other forces are apparently involved and selection seems most likely.

Estimates of the amount of admixture in various American black populations have been found to vary considerably (Workman et al., 1963; Blumberg and Hesser, 1971). Hemoglobin S and the G6PD deficiency give very high estimates of admixture. Since selection by malaria is not as great in the New World, these estimates are evidence of selection against these alleles. Some other loci such as the haptoglobins have high estimates, but the results are inconsistent. The significance of Blumberg and Hesser's (1971) analysis has been questioned by Mandarino and Cadien (1974), and Adams and Ward (1973) reanalyzed admixture rates for several American black populations using new estimates of the original African ancestral gene frequencies. They have found many impossible and inconsistent values from which they conclude that there is no evidence for selection. But the fact that outlandish estimates are obtained for some blood groups should not detract from the evidence of selection at other loci. The fitness differences are greatest among the genotypes for hemoglobin S and the G6PD deficiency, and it may well be that, given the number of generations, only selection of this magnitude can be detected. It is curious that the other locus associated with selection by malaria, the Duffy system, has generally been assumed to give the best estimates of admixture in black American populations (Reed, 1969). However, the very low estimate of admixture using the Fy^a allele for the Charleston blacks is probably due in part to the endemic vivax malaria that was present there.

Evidence from other hybrid populations is almost nonexistent. Morton et al. (1966) could find no evidence by segregation analysis for selection for 16 polymorphisms in the hybrid populations of Northeastern Brazil, with the exception of ABO blood group incompatibility. Crawford et al. (1976) used several methods to estimate admixture in the Tlaxcaltecs of Mexico, and for most of these methods the Duffy blood group gave significantly different estimates. This may be due to malaria selection, but their African frequencies seem to confuse Fy^b with Fy^o in Africans. In addition to Duffy, the Lewis and Diego blood groups and the Gc protein also gave admixture estimates that were significantly different from those expected.

Gene drift, like gene flow or admixture, has the same effect on all loci. Gene drift is measured by the variance in gene frequency among a set of isolates. This is expressed in terms of kinship or F-statistics with $F_{ST} = \sigma^2/q(1 - q)$, where σ^2 is the variance in gene frequency for any allele and q is its mean frequency. Some time ago I estimated the selective coefficients for the ABO blood groups in West Africa based on the greatly decreased variance for this locus compared to other blood groups (Livingstone, 1960), but over this great geographical range variation in selective coefficients would be expected and would give the same result. Cavalli-Sforza (1966) examined the variance

in blood groups in the Perma Valley and found little evidence for differences in variance. More recently, Lewontin and Krakauer (1973) devised a test to determine if the variances were significantly different. There has been much discussion of this test (Nei and Maruyama, 1975; Robertson, 1975; Ewens and Feldman, 1976), with a generally negative verdict on its validity.

Lewontin and Krakauer (1973) analyzed an early report on ten Yanomama villages and found significant heterogeneity among loci; but with a larger sample and more Yanomama villages included, there was no significant heterogeneity among loci (see also Neel and Ward, 1972; and Ward and Neel, 1976). Lewontin and Krakauer (1973) also show that there is significant heterogeneity among a worldwide sample of human populations for the Duffy and Rh alleles, which they state is due to selection. Although this may well be true, it is not because one allele has greater differences in fitness and thus the amount of gene drift is less. It is more likely that the equilibrium frequency varies more for one locus than another, and since the Duffy locus is associated with malaria, which has a restricted distribution in the tropics and subtropics, this may explain its greater variance.

In addition to Ward and Neel's (1976) analysis of the same Yanomama populations, Lewontin and Krakauer's test has been applied to the populations of Karkar Island (Boyce et al., 1978) and to those of the Markham River Valley, New Guinea (Wood, 1978), which were studied by Giles et al. (1966, 1970). In neither case was significant heterogeneity found, which seems to imply that drift and gene flow are responsible for the gene distributions among these isolates. However, Wood (1978) found no correlation between isolate size and heterozygosity, and this absence of association was also found among populations of the New Guinea Highlands by Wiesenfeld and Gajdusek (1976), which does not seem to support the explanation by gene drift.

Despite the promise that gene frequency analysis holds for detecting selection, few definite results that are generally agreed upon have been forthcoming (with the exception of the hemoglobin and G6PD loci) that have direct evidence for selection. This is undoubtedly due primarily to the smaller amount of selection occurring at other loci, but the type of selection is also likely to be involved. One of the implications of the sickle cell-malaria paradigm was the presence and importance of heterozygote advantage in balancing polymorphisms. This has been the type of selection assumed to occur in most theoretical analyses and for most human loci.

Two lines of evidence indicate that most human polymorphisms are not maintained by heterozygote advantage. First, for many loci there is too much polymorphism. Some *Drosophila* populations have over 20 and up to 40 alleles for some enzyme loci, and Lewontin et al. (1978) have shown that the set of heterotic fitness values that results in a stable equilibrium is vanishingly small and decreases rapidly with the number of alleles present. For the human species, the HLA loci (with an average of over ten alleles per locus) and the Gm and blood loci (with frequently as many alleles) are comparable and almost as difficult to maintain by heterosis. This is especially so for many small human isolates. Simulation of closed populations of up to 200 in size and 5% heterozygote advantage results in rapid extinction of most alleles within 100 generations. Second, the action of selection at many of the most polymorphic loci does not seem to accord with heterosis. For the HLA and blood group loci, particular antigens or alleles are associated with various diseases. For infectious diseases, the action of selection seems to be by "molecular mimicry" (Damian, 1964), which is difficult to reconcile with heterozygote advantage since heterozygotes that have an antigen in common with some parasite may be less affected than homozygotes but still be selected against to some degree. The action of most enzymes also does not seem to accord with heterozygote advantage but with a fitness intermediate between the homozygotes (Gillespie, 1978). However, it should be noted that for human enzyme loci, Harris et al. (1977) consider the greater amount of polymorphism among monomeric enzymes as evidence for a neutralist explanation, since these enzymes would probably have more neutral amino acid substitutions possible.

Heterozygote advantage is the simplest kind of selection that can result in balanced polymorphism, and the only kind with a single set of constant fitness values. It has

long been generally recognized that the way to balance polymorphisms was to develop models of changing fitness values. For predator and parasite interactions, frequency-dependent or density-dependent selection seem to be the most plausible alternatives (Clarke, 1975). Since predators tend to fix on the most common prey, they tend to take a greater proportion of them, which results in what has been called "apostatic selection." Similarly, parasites adapt to the antigens of the host, and this would appear to give a selective advantage to rare antigens. I have shown that frequency-dependent selection by disease could result in stable equilibria for the ABO blood groups, with fitness differences of about 5% (Livingstone, 1978). But it is very difficult to maintain ten alleles with only 5% selection, and simulation again shows that it is almost impossible to maintain this number of alleles in small isolates.

The simulation process has not been elaborated here, since simple genetic simulations are now well known and frequently are exercises in elementary computer courses. Briefly, a zygote is formed randomly by two gametes chosen from two previously chosen progenitors who have been randomly generated from the genotype frequencies of the adult generation. The zygote is then randomly selected out or stored according to its genotypic fitness, which can be constant or set proportional to the gene or genotypic frequency. The stored zygotes then form the next adult generation.

Most of the other models of variable selection can be described as either spatial or temporal, or even as both. Temporal models include both cyclical selection (Hoekstra, 1975), and random selection, in which the fitness values are randomly chosen each generation from a given distribution of values. Both models can lead to a balanced polymorphism, and Gillespie (1977, 1978) has shown that the random selection model can be made to fit enzyme activity, since heterozygotes have intermediate activity.

Spatial models have been of particular interest to zoologists, since many animal populations can inhabit a number of microhabitats (Powell and Taylor, 1979). Levene (1953) first developed a model with different selection in some parts of the range of a panmictic population, and since then there have been many elaborations of his model (Hedrick et al., 1976). Curiously, these models are very similar to those that have been proposed to demonstrate the possibility of group selection, (e.g., Wilson, 1975).

These models in many cases, however, do not seem appropriate for human populations. To have human groups occupying or "choosing" different niches and then coming together to mate seems unrealistic. However, the clines within the human species and even some of the polymorphisms within human populations are surely due to different selection operating in geographically separated populations and gene flow between them. But this is a somewhat different model. Nevertheless, spatial variation in fitness values is probably a major determinant of polymorphism and can result in stable equilibria (Felsenstein, 1976). The temporal variations in selection, whether cyclical or random, also do not seem in many ways to be appropriate for humans. There do not seem to be major selective forces that "cycle" over a period of several generations, and the causes of selection, differential mortality, and fertility do not seem to be randomly distributed from generation to generation.

Human existence is more complicated, and with our long life-span there are many selective episodes throughout life. The preceding models have assigned one fitness to a genotype for its lifetime, but we are continually being challenged. The history of life in general has been characterized as "long periods of boredom and short periods of terror" (Gould, 1977), and from the standpoint of selection this seems to be appropriate for a life history. For example, the sickle cell trait carrier and normal have the same fitness most of the time in Africa; it is only during that first lethal struggle with the malaria parasite at some time between 6 months and 3 years of age that the difference occurs. Thus, for any locus there can be many episodes of selection throughout life for which the genotypes have different fitnesses. This is a form of temporal selection and can result in a balanced polymorphism with a variety of fitness values. For the blood groups and other loci that have been associated with a great number of diseases, there would be a number of such episodes. I have run simulations with three episodes, each one favoring a different allele, and all three can be maintained in populations as small as 200. However, ten episodes do not seem to be enough to maintain ten alleles in populations of this size.

In conclusion, we now have a great many models that can maintain our large amount of polymorphism, although we cannot measure the fitness differences. There is little doubt that many amino acid substitutions appear to be neutral or quasi-neutral, but neutral models do not seem to fit the distributions of most loci in the human species, especially those loci with considerable geographical variation.

LITERATURE CITED

- Adams, J., and R.H. Ward (1973) Admixture studies and the detection of selection. *Science*, *180*:1137-1143.
- Allison, A.C. (1954) Protection afforded by sickle-cell trait against subtertian malarial infection. *Br. Med. J.*, *1*:290-294.
- Ananthkrishnan, R., and H. Walter (1972) Some notes on the geographical distribution of the human red cell acid phosphatase phenotypes. *Humangenetik*, *15*:177-181.
- Andermann, E., C.R. Scriver, R. Gold, L. Wolfe, G. Patry, R. Lafontaine, G. Geoffroy, and F. Andermann (1975) Tay Sachs screening in a French Canadian deme: Suggestive evidence for polymorphism. *Am. J. Hum. Genet.*, *27*:12A.
- Beals, K.L., and A.J. Kelso (1975) Genetic variation and cultural evolution. *Am. Anthropol.*, *77*:566-579.
- Beckman, G., L. Beckman, B. Cedergren, and S. Liden (1977) Red cell enzyme types in psoriasis. *Hum. Hered.*, *27*:408-411.
- Beet, E.A. (1946) Sickle cell disease in the Balovale district of Northern Rhodesia. *E. Afr. Med. J.*, *23*:75-86.
- Beet, E.A. (1947) Sickle cell disease in Northern Rhodesia. *E. Afr. Med. J.*, *24*:212-222.
- Bianco, I. (1949) Ricerche sulle resistenze osmotiche delle emazie nei malarici. *Riv. Malariol.*, *28*:331-342.
- Bienzle, U., O. Aveni, A.O. Lucas, and L. Luzzatto (1972) Glucose-6-phosphate dehydrogenase and malaria. *Lancet*, *1*:107-110.
- Blumberg, B.S., and J.E. Hesser (1971) Loci differentially affected by selection in two American Black populations. *Proc. Natl. Acad. Sci. U.S.A.*, *68*:2554-2558.
- Bodmer, W.F., and J.G. Bodmer (1978) Evolution and function of the HLA system. *Br. Med. Bull.*, *34*:309-316.
- Bois, E., J. Feingold, F. Demezis, Y. Runavot, M. Jehanne, and L. Tondic (1978) Cluster of cystic fibrosis cases in a limited area of Brittany (France). *Clin. Genet.*, *14*:73-76.
- Boyce, A.J., G.A. Harrison, C.M. Platt, R.W. Hornabrook, S. Serjeantson, R.K. Kirk, and P.B. Booth (1978) Migration and genetic diversity in an island population: Karkar, Papua New Guinea. *Proc. R. Soc. Lond. B*, *202*:269-295.
- Boyd, W.C. (1950) *Genetics and the Races of Man*. Little, Brown, Boston.
- Boyd, W.C. (1963) *Genetics and the human race*. *Science*, *140*:1057-1064.
- Boyer, S.H., A.F. Scott, L.M. Kunkel, and K.D. Smith (1978) The proportion of all point mutations which are unacceptable: An estimate based on hemoglobin amino acid and nucleotide sequences. *Can. J. Genet. Cytol.*, *20*:111-137.
- Brachtel, R., H. Walter, W. Beck, and M. Hilling (1979) Associations between atopic diseases and the polymorphic systems ABO, Kidd, Inv, and red cell acid phosphatase. *Hum. Genet.*, *49*:337-348.
- Brain, P. (1952a) Sickle-cell anaemia in Africa. *Br. Med. J.*, *2*:880.
- Brain P. (1952b) The sickle-cell trait. *S. Afr. Med. J.*, *26*:925-928.
- Cavalli-Sforza, L.L. (1966) Population structure in human evolution. *Proc. R. Soc. Lond. B*, *164*:362-379.
- Cavalli-Sforza, L.L., and W.F. Bodmer (1971) *The Genetics of Human Populations*. W.H. Freeman, San Francisco.
- Cavalli-Sforza, L.L., and A.W.F. Edwards (1963) Analysis of human evolution. In: *Genetics Today*. S.J. Geerts, ed. (Proc. XI Int. Cong. Genet., 3:923-933), Pergamon Press, New York.
- Chakravarti, A., and R. Chakraborty (1978) Elevated frequency of Tay-Sachs disease among Ashkenazic Jews unlikely by genetic drift alone. *Am. J. Hum. Genet.*, *30*:256-261.
- Chase, G.A., and V.A. McKusick (1972) Founder effect in Tay-Sachs disease. *Am. J. Hum. Genet.*, *24*:339-340.
- Chern, C.J., and E. Beutler (1976) Biochemical and electrophoretic studies of erythrocyte pyridoxine kinase in white and black Americans. *Am. J. Hum. Genet.*, *28*:9-17.
- Clarke, B. (1975) Frequency-dependent and density-dependent natural selection. In: *The Role of Natural Selection in Human Evolution*. F.M. Salzano, ed. American Elsevier, New York, pp. 187-202.
- Coon, C.S. (1962) *The Origin of Races*. Alfred A. Knopf, New York.
- Crawford, M.H. (1973) The use of genetic markers of the blood in the study of the evolution of human populations. In: *Methods and Theories of Anthropological Genetics*. M.H. Crawford and P.L. Workman, eds. University of New Mexico Press, Albuquerque, pp. 19-38.
- Crawford, M.H., P.L. Workman, C. McLean, and F.C. Lees (1976) Admixture estimates and selection in Tlaxcala. In: *The Tlaxcaltecs: Prehistory, Demography, Morphology and Genetics*. M.H. Crawford, ed. University of Kansas Publications in Anthropology, Number 7, Lawrence.
- Crawford, M. d'A. (1975) Frequency of cystic-fibrosis gene. *Lancet*, *1*:167.
- Daigher, S.P., M.S. Schanfield, and L.L. Cavalli-Sforza (1975) Group-specific component (Gc) proteins bind vitamin D and 25-hydroxyvitamin D. *Proc. Natl. Acad. Sci. USA*, *72*:2076-2080.
- Damian, R.T. (1964) Molecular mimicry: Antigen sharing by parasite and host and its consequences. *Am. Nat.*, *98*:129-149.
- Das, B.M., R. Deka, and G. Flatz (1975) Predominance of the haemoglobin E gene in a Mongoloid population in Assam (India). *Humangenetik*, *30*:187-191.
- Dobzhansky, T. (1950) The genetic nature of differences among men. In: *Evolutionary Thought in America*. S. Persons, ed. Yale University Press, New Haven, pp. 86-155.
- Dozy, A.M., Y.W. Kan, S.H. Embury, W.C. Mentzer, and W.C. Wang (1979) α -globin gene organization in blacks precludes the severe form of α -thalassaemia. *Nature*, *280*:605-607.
- Eaton, J.W., J.R. Eckman, E. Berger, and H.S. Jacob (1976) Suppression of malaria infection by oxidant-

- sensitive host erythrocytes. *Nature*, 264:758-760.
- Edwards, A.W.F., and L.L. Cavalli-Sforza (1964) Reconstruction of evolutionary trees. In: Phenetic and Phylogenetic Classification. V.H. Heywood and J. McNeill, eds. The Systematics Association, Publication Number 6, London, pp. 67-76.
- Ewens, W.J., and M.W. Feldman (1976) The theoretical assessment of selective neutrality. In: Population Genetics and Ecology. S. Karlin and E. Nevo, eds. Academic Press, New York, pp. 303-337.
- Felsenstein, J. (1976) The theoretical population genetics of variable selection and migration. *Annu. Rev. Genet.*, 10:253-280.
- Fitch, W.M., and J.V. Neel (1969) The phylogenetic relationships of some Indian tribes of Central and South America. *Am. J. Hum. Genet.*, 21:384-397.
- Fraikor, A.L. (1977) Tay-Sachs disease: Genetic drift among the Ashkenazim Jews. *Soc. Biol.*, 24:117-134.
- Friedman, M.J. (1978) Erythrocytic mechanism of sickle cell resistance to malaria. *Proc. Natl. Acad. Sci. USA*, 75:1994-1997.
- Friedman, M.J. (1979) Oxidant damage mediates variant red cell resistance to malaria. *Nature*, 280:245-247.
- Giblett, E.R. (1969) Genetic Markers in Human Blood. Blackwell Scientific Publications, Oxford.
- Giblett, E.R. (1977) Genetic polymorphisms in human blood. *Annu. Rev. Genet.*, 11:13-28.
- Giles, E., E. Ogan, R.J. Walsh, and M.A. Bradley (1966) Blood group genetics of natives of the Morobe District and Bougainville, Territory of New Guinea. *Archaeol. Phys. Anthropol. Oceania*, 1:135-154.
- Giles, E., S. Wyber, and R.J. Walsh (1970) Micro-evolution in New Guinea: Additional evidence for genetic drift. *Archaeol. Phys. Anthropol. Oceania*, 5:60-72.
- Gillespie, J.H. (1977) A general model to account for enzyme variation in natural populations. III. Multiple alleles. *Evolution*, 31:85-90.
- Gillespie, J.H. (1978) A general model to account for enzyme variation in natural populations. V. The SAS-CFF model. *Theoret. Pop. Biol.*, 14:1-45.
- Gould, S.J. (1977) Evolution's erratic pace. *Nat. Hist.*, 86:12, 14, 16.
- Haigh, J., and J. Maynard-Smith (1972) Population size and protein variation in man. *Genet. Res.*, 19:73-89.
- Haldane, J.B.S. (1949) Disease and evolution. *La Ricerca Scientifica*, 19(Suppl):1-11.
- Harpending, H. (1974) Genetic structure of small populations. *Annu. Rev. Anthropol.*, 3:229-243.
- Harris, H., D.A. Hopkinson, and Y.H. Edwards (1977) Polymorphism and the subunit structure of enzymes: A contribution to the neutralist-selectionist controversy. *Proc. Natl. Acad. Sci. USA*, 74:698-701.
- Harrison, G.A., A.J. Boyce, R.W. Hornabrook, S. Serjeantson, and W.J. Craig (1976) Evidence for an association between ABO blood group and goitre. *Hum. Genet.*, 32:335-337.
- Hedrick, P.W., M.E. Ginevan, and E.P. Ewing (1976) Genetic polymorphism in heterogeneous environments. *Annu. Rev. Ecol. Syst.*, 7:1-32.
- Hedrick, P.W., and E. Murray (1978) Average heterozygosity revisited. *Am. J. Hum. Genet.*, 30:377-382.
- Hiernaux, J., and A. Froment (1976) The correlations between anthropobiological and climatic variables in Sub-Saharan Africa: Revised estimates. *Hum. Biol.*, 48:757-767.
- Hitzeroth, H.W., H. Walter, and M. Hilling (1978) Genetic markers and leprosy in South African Negroes. *South Afr. Med. J.*, 54:653-657.
- Hoekstra, R.F. (1975) A deterministic model of cyclical selection. *Genet. Res.*, 25:1-15.
- Huxley, J., E. Mayr, H. Osmond, and A. Hoffer (1964) Schizophrenia as a genetic morphism. *Nature*, 204:220-221.
- Kelly, T.E., G.A. Chase, M.M. Kaback, K.K. Komur, and V.A. McKusick (1975) Tay-Sachs disease: High gene frequency in a non-Jewish population. *Am. J. Hum. Genet.*, 27:287-291.
- Kimura, M. (1968) Evolutionary rate at the molecular level. *Nature*, 217:624-626.
- King, J.L., and T.H. Jukes (1969) Non-Darwinian evolution. *Science*, 164:788-798.
- Knudsen, A.G. (1979) Our load of mutations and its burden of disease. *Am. J. Hum. Genet.*, 31:401-413.
- Levene, H. (1953) Genetic equilibrium when more than one ecological niche is available. *Am. Nat.*, 87:331-333.
- Lewontin, R.C. (1974) *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.
- Lewontin, R.C., L.R. Ginzburg, and S.D. Tuljapurkar (1978) Heterosis as an explanation for large amounts of genic polymorphism. *Genetics*, 88:149-170.
- Lewontin, R.C., and J. Krakauer (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, 74:175-195.
- Livingstone, F.B. (1960) Natural selection, disease, and ongoing human evolution, as illustrated by the ABO blood groups. *Hum. Biol.*, 32:17-27.
- Livingstone, F.B. (1967) *Abnormal Hemoglobins in Human Populations*. Aldine, Chicago.
- Livingstone, F.B. (1969) The founder effect and deleterious genes. *Am. J. Phys. Anthropol.*, 30:55-60.
- Livingstone, F.B. (1971) Malaria and human polymorphisms. *Annu. Rev. Genet.*, 5:33-64.
- Livingstone, F.B. (1973) Data on the abnormal hemoglobins and glucose-6-phosphate dehydrogenase deficiency in human populations, 1967-1973. Museum of Anthropology, University of Michigan, Technical Report Number 3, Ann Arbor.
- Livingstone, F.B. (1978) Frequency-dependent selection and the ABO blood groups. In: *Evolutionary Models and Studies in Human Diversity*. R.J. Meier, C.M. Otten and F. Abel-Hameed, eds. Mouton, New York, pp. 127-139.
- Livingstone, F.B. (in press) Natural selection and random variation in human populations. In: *Current Developments in Anthropological Genetics*, vol. 1 M.H. Crawford, ed. Plenum, New York.
- Lucarelli, P., R. Scacchi, R.M. Corbo, G. Laconi, A. Nasi, and R. Palmirino (1979) Placental alkaline phosphatase: population studies in Sardinia and data on the anthropological value of this genetic marker. *Am. J. Phys. Anthropol.*, 50:605-610.
- Lucarelli, R., R. Scacchi, R.M. Corbo, P. Palmirino, M. Orsini, L. Campea, E. Carapella, and R. Pascone (1978) Genetic polymorphisms in juvenile-onset diabetes. *Hum. Hered.*, 28:89-95.

- Luzzatto, L., and U. Bienle (1979) The malaria/G-6-P.D. hypothesis. *Lancet*, 1:1183-1184.
- Mandarino, L., and J.D. Cadien (1974) Use of ranked migration estimates for detecting natural selection. *Am. J. Hum. Genet.*, 26:108-110.
- Martin, S.K., L.H. Miller, D. Alling, C.V. Okoye, G.J.F. Esan, B.O. Osunkoya, and M. Deane (1979) Severe malaria and glucose-6-phosphate-dehydrogenase deficiency: a reappraisal of the malaria/G-6-P.D. hypothesis. *Lancet*, 1:524-526.
- Martin, S.K., L.H. Miller, J.A. Kark, C.U. Hicks, M.J. Haut, V.C. Okoye, and G.J.F. Esan (1978) Low erythrocyte pyridoxal-kinase activity in blacks: Its possible relationship to falciparum malaria. *Lancet*, 1:466-468.
- McCullough, J.M., and E. Giles (1970) Human cerumen types in Mexico and New Guinea: A humidity related polymorphism in "Mongoloid" peoples. *Nature*, 226:460-462.
- McKusick, V.A. (1978) Mendelian Inheritance in Man, 5th edition. John Hopkins Press, Baltimore.
- Miller, L.H., S.J. Mason, D.F. Clyde, and M.H. McGinnis (1976) The resistance factor to *Plasmodium vivax* in blacks. The Duffy blood group genotype, FyFy. *N. Engl. J. Med.*, 295:302-304.
- Miller, L.H., M.H. McGinnis, P.V. Holland, and P. Sigmon (1978) The Duffy blood group phenotype in American blacks infected with *Plasmodium vivax* in Vietnam. *Am. J. Trop. Med. Hyg.*, 27:1069-1072.
- Modiano, G., G. Battistuzzi, G.J.F. Esan, U. Testa, and L. Luzzatto (1979) Genetic heterogeneity of "normal" human erythrocyte glucose-6-phosphate dehydrogenase: An isoelectrophoretic polymorphism. *Proc. Natl. Acad. Sci. USA*, 76:852-856.
- Morton, N.E. (1972) Population genetics. In: *Human Genetics (Proc. IV Int. Cong. Hum. Genet.)* J. de Grouchy, F.J.G. Ebling and I.W. Henderson, eds. Excerpta Medica, Amsterdam, pp. 131-132.
- Morton, N.E., H. Krieger, and M.P. Mi (1966) Natural selection on polymorphisms in northeastern Brazil. *Am. J. Hum. Genet.*, 18:153-171.
- Motulsky, A.G. (1975) G6PD and abnormal hemoglobin polymorphisms—Evidence regarding malarial selection. In: *The Role of Natural Selection in Human Evolution*. F.M. Salzano, ed. American Elsevier, New York, pp. 271-294.
- Mourant, A.E., A.C. Kopec, and K. Domaniewska-Sobczak (1976a) The distribution of the human blood groups and other polymorphisms. Oxford University Press, London.
- Mourant, A.E., A.C. Kopu, and K. Domaniewska-Sobczak (1978) *Blood Groups and Diseases*. Oxford University Press, Oxford.
- Mourant, A.E., D. Tills, and K. Domaniewska-Sobczak (1976b) Sunshine and the geographical distribution of the alleles of the Gc system of plasma proteins. *Hum. Genet.*, 33:307-314.
- Myrionthopoulos, N.C. (1972) Population dynamics of Tay-Sachs disease. II. What confers the selective advantage upon the Jewish heterozygote? In: *Sphingolipids, Sphingolipidosis, and Allied Disorders*. Advances in Experimental Medicine and Biology. B.W. Volk and S.M. Aronson, eds. Plenum, New York, pp. 561-569.
- Myrionthopoulos, N.C., A.F. Naylor, and S.M. Aronson (1972) Founder effect in Tay-Sachs disease unlikely. *Am. J. Hum. Genet.*, 24:341-342.
- Neel, J.V. (1978) Rare variants, private polymorphisms, and locus heterozygosity in Amerindian populations. *Am. J. Hum. Genet.*, 30:465-490.
- Neel, J.V., H.B. Hamilton, T.Y. Kobara, and K. Ozaki (1963) The uneven distribution in Japan of carriers of the rare recessive gene causing acatalasemia and the implications for studies on inbreeding effects. In: *The Genetics of Migrant and Isolate Populations*. E. Goldschmidt, ed. Williams and Wilkins, Baltimore, pp. 176-183.
- Neel, J.V., and R.H. Ward (1972) The genetic structure of a tribal population, the Yanomama Indians. VI. Analysis by F-statistics (including a comparison with the Makiritare and Xavante). *Genetics*, 72:639-666.
- Nei, M., and T. Maruyama (1975) Lewontin-Krakauer test for neutral genes. *Genetics*, 80:395.
- Nei, M., and A.K. Roychoudhury (1974) Genic variation within and between the three major races of man, Caucasoids, Negroids, and Mongoloids. *Am. J. Hum. Genet.*, 26:421-443.
- Pasvol, G., D.J. Weatherall, and R.J.M. Wilson (1977) Effects of foetal haemoglobin on susceptibility of red cells to *Plasmodium falciparum*. *Nature*, 270:171-173.
- Pasvol, G., D.J. Weatherall, and R.J.M. Wilson (1978) Cellular mechanism for the protective effect of haemoglobin S against *P. falciparum* malaria. *Nature*, 274:701-703.
- Piazza, A., and C. Viganotti (1973) Evolutionary trees and HL-A polymorphism. In: *Histocompatibility Testing 1972*. J. Dausset and J. Colombani, eds. Munksgaard, Copenhagen, pp. 731-738.
- Powell, J.R., and C.E. Taylor (1979) Genetic variation in ecologically diverse environments. *Am. Sci.*, 67:590-596.
- Raper, A.B. (1949) The incidence of sickle cell anemia. *E. Afr. Med. J.*, 26:281-282.
- Reed, T.E. (1969) Caucasian genes in American Negroes. *Science*, 165:762-768.
- Reed, T.E. (1975) Selection and the blood group polymorphisms. In: *The Role of Natural Selection in Human Evolution*. F.M. Salzano, ed. American Elsevier, New York, pp. 231-246.
- Robertson, A. (1975) Gene frequency distributions as a test of selective neutrality. *Genetics*, 81:775-785.
- Roth, E.F., M. Friedman, Y. Ueda, I. Tellez, W. Trager, and R.L. Nagel (1978) Sickling rates of human AS red cells infected in vitro with *Plasmodium falciparum* malaria. *Science*, 202:650-652.
- Sarich, V.M. (1977) Rates, sample sizes, and the neutrality hypothesis for electrophoresis in evolutionary studies. *Nature*, 265:24-28.
- Schell, L.M. S.S. Agarwal, B.S. Blumberg, H. Levy, P.H. Bennett, W.S. Laughlin, and J.P. Martin (1978) Distribution of albumin variants Naskapi and Mexico among Aleuts, Frobisher Bay Eskimos, and Micmac, Naskapi, Mohawk, Omaha and Apache Indians. *Am. J. Phys. Anthropol.*, 49:111-118.
- Serjeantson, S., K. Bryson, D. Amato, and D. Babona (1977) Malaria and hereditary ovalocytosis. *Hum. Genet.*, 37:161-167.
- Spencer, H.C., L.H. Miller, W.E. Collins, C. Knud-Hansen, M.H. McGinnis, T. Shiroishi, R.A. Lobos, and R.A.

- Feldman (1978) The Duffy blood group and resistance to *Plasmodium vivax* in Honduras. *Am. J. Trop. Med. Hyg.*, 27:664-670.
- Spielman, R.S., E.C. Migliazza, J.V. Neel, H. Gershowitz, and R. Torres de Arauz (1979) The evolutionary relationships of two populations: A study of the Guaymi and the Yanomama. *Current Anthropol.*, 20:377-388.
- Stuart, A.B., and M.G. Burdon (1974) Frequency of cystic-fibrosis gene. *Lancet*, 2:1521.
- Super, M. (1975) Cystic fibrosis in the South West Afrikaner. *S. Afr. Med. J.*, 49:818-820.
- Szathmary, E.J.E., and N.S. Ossenberg (1978) Are the biological differences between North American Indians and Eskimos truly profound? *Current Anthropol.*, 19:673-701.
- Vezzoso, B. (1946) Influenza della malaria sulla mortalita infantile per anemia con speciale riguardo al morbo di Cooley. *Riv. Malariol.*, 25:61-77.
- Vogel, F. (1975) ABO blood groups, the HL-A system and diseases. In: *The Role of Natural Selection in Human Evolution*. F.M. Salzano, ed. American Elsevier, New York, pp. 247-269.
- Wagener, D., L.L. Cavalli-Sforza, and R. Barakat (1978) Ethnic variation of genetic disease: Roles of drift for recessive lethal genes. *Am. J. Hum. Genet.*, 30:262-270.
- Walter, H., and M. Bajatzadeh (1971) Investigations on the geographical variability of the human transferrins. *Humangenetik*, 12:267-274.
- Ward, R.H., and J.V. Neel (1976) The genetic structure of a tribal population, the Yanomama Indians. XIV. Clines and their interpretation. *Genetics*, 82:103-121.
- Weatherall, D.J., and J.B. Clegg (1979) Recent developments in the molecular genetics of human hemoglobin. *Cell*, 16:467-479.
- Welch, S.G., I.A. McGregor, and K. Williams (1977) The Duffy blood group and malaria prevalence in Gambian West Africans. *Trans. Roy. Soc. Trop. Med. Hyg.*, 71:295-297.
- Wiensfeld, S.L., and D.C. Gajdusek (1976) Genetic structure and heterozygosity in the kuru region, eastern highlands of New Guinea. *Am. J. Phys. Anthropol.*, 45:177-190.
- Wilson, D.S. (1975) A theory of group selection. *Proc. Natl. Acad. Sci. USA*, 72:143-146.
- Winsor, E.J.T., and J.P. Welch (1978) Genetic and demographic aspects of Nova Scotia Niemann-Pick disease (Type D). *Am. J. Hum. Genet.*, 30:530-538.
- Wood, J.W. (1978) Population structure and genetic heterogeneity in the upper Markham Valley of New Guinea. *Am. J. Phys. Anthropol.*, 48:463-470.
- Workman, P.L., B.S. Blumberg, and A.J. Cooper (1963) Selection, gene migration and polymorphic stability in a U.S. white and Negro population. *Am. J. Hum. Genet.*, 15:429-437.
- Wright, S. (1951) The genetical structure of populations. *Ann. Eugen.*, 15:322-354.
- Yamazaki, T., and T. Maruyama (1972) Evidence for the neutral hypothesis of protein polymorphism. *Science*, 178:56-57.
- Yamazaki, T., and T. Maruyama (1974) Evidence that enzyme polymorphisms are selectively neutral, but blood group polymorphisms are not. *Science*, 183:1091-1092.