THE DISTRIBUTIONS OF THE ABNORMAL HEMOGLOBIN GENES AND THEIR SIGNIFICANCE FOR HUMAN EVOLUTION

FRANK B. LIVINGSTONE

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

One of the most spectacular advances in the last 10 years in our knowledge of human genetic diversity has been the discovery of the abnormal hemoglobin genes. In this short space of time the distributions of these genes in the world's populations have become one of the best known genetic systems for any animal species. In fact, the great mass of data makes it difficult to summarize our knowledge of these genes in a single paper, and the heterogeneity of these data emphasizes once again the great genetic variability which exists among the populations of the human species. Although we know more about the hemoglobins, this is but one of the many genetic systems which the increasing precision of biochemical techniques has brought to our attention. The haptoglobins, transferrins, and others yet to be discovered will increase our knowledge of human genetic diversity still further.

Another significant advance in recent years, which preceded the discovery of the hemoglobins, was the development of the modern genetic theory of evolution. This synthesis was a theoretical advance as much as a factual one and was developed primarily from data on animals other than man. It has been applied to some extent to the blood group genes, but because we know so much more about the factors of evolution as regards the abnormal hemoglobin genes, this system is the first one for the human species for which we can discuss the factors of evolution and their interrelationships. Since, in addition, we also know the environmental circumstances which determine the direction of natural selection, the principal factor of evolution, we can effectively evaluate the role of culture as one of the environmental determinants of human evolution at the hemoglobin locus.

While the populational aspects of the abnormal hemoglobin genes have been of great interest to anthropologists and geneticists, biochemists have also made significant advances in the chemical structure of hemoglobin so that hemoglobin is one of the best known complex proteins. The chemical structure of this molecule has been directly related to gene action as no other has. We now know that genes are composed of DNA, which, through RNA, directs the manufacture of proteins or polypeptide chains, but for the hemoglobin molecule we know the exact changes in the molecule which are associated with the gene differences. (The biochemical advances are discussed by Ingram, 1963.)

As the specific changes which genes produce have been discovered, it has been realized that similar phenotypic characteristics, whether based on observation with the naked eye or biochemical tests, can result from very different genetic structures. For example, one can talk about the gene for albinism and even calculate the mutation rate to this gene, but we now realize that literally dozens of different genetic changes may result in phenotypically similar albinism. This specificity of genes also has implications for measuring gene flow between populations. Many of the characteristics which have been used to claim genetic relationships between populations may well be due to very different genes. The abnormal hemoglobins, as the first system where we know the specific chemical changes involved, have made it possible to discuss the basic concepts of mutation, gene flow, and natural selection in a much more realistic way.

In 1953 after the development of paper electrophoresis made mass surveys for the abnormal hemoglobins quite easy, an international committee was convened to

establish some orderly process for the naming of the new hemoglobins which were rapidly being discovered. This committee established the letters A, F, M, and S, for normal, fetal, met-, and sickle cell hemoglobins which had been known for some time and decided that further hemoglobins would be assigned letters in order of discovery, beginning with the letter C, and working up the alphabet. At present we are up to the letter R, but further developments have rendered this system rather unworkable as originally set up. These developments were the perfection of techniques to determine the complete chemical structure of hemoglobin. They have resulted in the discovery that hemoglobins which were being called by the same letter since they had the same electrophoretic properties were actually quite different in structure. Hence, another conference has resulted in the use of subscripts to the letters to distinguish these similar hemoglobins. These subscripts designate the place of discovery so that we now have, for example, hemoglobin $G_{San\ Jose}$, hemoglobin $G_{Bristol}$, and hemoglobin $G_{Philadelphia}$.

With a few exceptions which are more complicated genetically, all of these newly discovered abnormal hemoglobins are due to the presence of a single gene, or in other words differ from normal hemoglobin in one part of their structure. When heterozygous, these genes are for the most part benign, while homozygosity or simultaneous heterozygosity for two of them result in conditions of variable severity ranging from lethality for the sickle cell homozygote in the absence of modern medicine to a slight anemia as in homozygosity for hemoglobin C. For many of the rare hemoglobins no homozygotes have been detected, but it would appear that these conditions may well be more like hemoglobin C in severity than the sickle cell gene.

Although the hemoglobin genes may still hold some surprises for us, the world distribution of these genes can be mapped in considerable detail. For purposes of further discussion the hemoglobin genes will be divided into three groups: (1) those which attain appreciable frequencies or more than 15% heterozygotes in some populations, (2) those found in intermediate frequencies or more than 1% but less than 15% heterozygotes, and (3) those found only occasionally or in less than 1% of the population. The letter designations of the hemoglobins will be used although it should be remembered that any one of these letters may indicate more than one specific gene. With this classification hemoglobins S, C, E, and thalassemia would fall into group 1; hemoglobins D, K, N, and O in group 2; and the rest in group 3. Thalassemia is a potpourri and undoubtedly contains many different "genes." This is because thalassemia is still primarily diagnosed on clinical and morphological features, and many cases which appear similar in these characteristics have been shown to have different chemical and genetic bases. Some thalassemias are chromosomal duplications due to unequal crossing over, some are probably specific structural changes in the hemoglobin molecule as are the other abnormal hemoglobins, and others may result from operator or regulator genes, whatever that means. Nevertheless, some populations have high frequencies of all thalassemias combined so that this system will be considered in the same category as hemoglobins S, C, and E, which are more specific entities.

Some 25 different rare hemoglobins have been discovered so far, although in order to discover these, extremely large samples must be collected. Many of these rare hemoglobins have been discovered in populations in the tropical and subtropical regions of the Old World which have high frequencies of abnormal hemoglobin genes, but the rare hemoglobins have also been discovered in populations outside this area. Studies from Japan indicate that these random mutants occur in about one of every 2,000 individuals examined (Shibata, 1961). We have talked about mutation as a random, recurrent process and we now have this excellent example. Furthermore, most mutations were thought to be deleterious and this appears to be true of the

hemoglobins. Thus, the low frequency of abnormal hemoglobins in the Japanese, English, or Swedish populations results from a balance between mutation to these genes and selection against them. This type of selection which keeps the gene frequencies where they are has been called stabilizing selection, and the hemoglobin loci in these populations are examples of the classical theory of genetic diversity which views this diversity as a balance between mutation away from the normal genotype and natural selection weeding out these deleterious mutants. Evolution or gene frequency change has not occurred here.

The classical view of evolutionary change is usually stated as "every once in awhile a favorable mutation occurs which then replaces the older norm." For the abnormal hemoglobins this statement is wrong or at least it seems to me to be the wrong way to characterize their evolution. It is rather backward and another example of genetic determinism which is endemic in Western thought. It is important to remember that mutation is a recurrent, continuous process, and our hominid and pongid ancestors were continually mutating and producing abnormal hemoglobins. But little evolution occurred; so little, in fact, that based on peptide fingerprinting, the hemoglobin of most of mankind is more similar to that of the chimpanzee and gorilla than it is to their fellow-man, including in some cases immediate family. The cause of evolutionary change in the human hemoglobin genes is not the occurrence of favorable mutants, but changes in the amount and direction of natural selection operating on these genes, and these changes in natural selection were the result of the development of environmental circumstances in which heterozygotes for some of these abnormal hemoglobin genes possessed a selective advantage. It was only then that these genes became an important part of some human gene pools.

While all populations which have been intensively investigated have low frequencies of some abnormal hemoglobin genes, it is only populations which have been resident for a considerable length of time in the tropical and subtropical regions of the Old World which have intermediate or appreciable frequencies of abnormal hemoglobins. However, there are populations in this area, such as the Australian aborigines, some Papuans, Polynesians, and Micronesians in the Asian tropics and the Bushmen, Ethiopians, and some Nilotic peoples in the African tropics, that do not have high frequencies of any abnormal hemoglobin. With the exception of thalassemia, the hemoglobins which attain high frequencies in the tropics, hemoglobins S, C, and E, are rarely found outside their areas of high frequencies except in migrant populations. Thus, they do not seem to be frequent mutants of the hemoglobin loci, but relatively rare ones.

Although thalassemia is a genetically heterogeneous entity, hemoglobins S, C, and E are quite specific, and all differ from normal hemoglobin by only a single amino acid out of the string of about 300 amino acids which comprise the identical halves of a hemoglobin molecule. Since an abnormal hemoglobin mutation presumably could occur on most of these 300 amino acids, it is obvious that the mutation rates to the hemoglobin S, C, or E gene are much less than the mutation rate to thalassemia which includes many different changes.

Since hemoglobins S, C, and E appear to be relatively rare mutants and their distributions are contiguous and somewhat restricted, gene flow appears to have been a more important factor in determining their distributions than in the case of thalassemia which is a common mutant. Of course, to explain similar frequencies of abnormal hemoglobins in any two populations, one has to weigh the possibility of gene flow with the possibility of a separate mutation becoming established. But this is possible in the case of these three hemoglobins since all frequencies seem to be of the same specific gene mutation in each case, while for thalassemia the recognition of gene flow is difficult. It should be noted that all that sickles is not hemoglobin S;

there are other hemoglobins which also produce the sickling phenomenon, but the distributions which will be discussed are based for the most part on electrophoresis tests and not on sickle cell tests with sodium metabisulfite.

All four of the abnormal hemoglobins which are found in high frequencies have some deleterious effects associated with homozygosity. In fact, it is rather surprising that thalassemia and hemoglobin S, the most widespread and most frequent of the four, have the most serious effects. Although some individuals who appear to be homozygous for either hemoglobin S or thalassemia have survived and reproduced, population studies indicate that in the relatively primitive medical conditions of the areas of Africa and Europe where high frequencies of these genes are found, the survival and reproduction of homozygotes for these genes is practically nil. Such selection against the sickle cell and thalassemia genes thus raises the question as to why there are high frequencies of these genes in some populations.

Neel (1951) first suggested the two possible answers, mutation or heterozygote superiority, and there were suggestions that malaria may be the factor which confers an increased fitness on the heterozygote. But Allison (1954a, 1954b, 1954c) first grasped the problem and attempted to answer it. His investigations, although providing strong evidence of an association between the sickle cell gene and malaria, were not conclusive; but since then investigations by Raper (1956), the Lambotte-Legrands (1957), Delbrouck (1958), and Firschein (1961) have demonstrated conclusively that sickle cell trait carriers have a selective advantage in an area with holoendemic malaria which is due primarily to *Plasmodium falciparum*. Although disbelief in the malaria hypothesis is still occasionally expressed, I think this work is an excellent example of progress in science. The question, why high frequencies of the sickle cell gene, is a legitimate one and must have an answer. In terms of the modern theory of population genetics there may be several

answers, but I have never seen an alternative to the malaria hypothesis seriously advanced. And in terms of the total array of data we now possess on this gene, there is no other answer. Assuming that this is a balanced polymorphism, the factor which is balancing the polymorphism in many populations of Africa, Greece, and India would have to eliminate about 15% of the normal homozygotes each generation or about 11% of the total population. Malaria is the only cause of death which comes anywhere near eliminating this many zygotes. Thus, in terms of our theories, this question must have an answer and this is the only possible one.

Although the relationship between sickling and malaria has much evidence in its favor, practically no work has been done on the selective factors of the other abnormal hemoglobin genes. In Greece (Choremis, Fessas, Kattamis, Stamatoyannopoulos, Zannos-Mariolea, Karaklis, and Belios, 1963), Sardinia (Ceppellini, 1957), and New Guinea (Curtain, Kidson, Gajdusek, and Gorman, 1962) the frequency of thalassemia has been shown to be correlated with the amount of malaria, and there is some evidence from Ghana (Thompson, 1962) that hemoglobin C carriers may have some resistance to malaria infections. By analogy with hemoglobin S, it has been generally assumed that malaria is the major factor involved in the other abnormal hemoglobins, but since the differences in fitness which we are attempting to detect are so much smaller for the abnormal hemoglobins, C, E, and, in most areas, thalassemia, the task is much more difficult than for the sickle cell gene. Furthermore, malaria is rapidly being controlled as a lethal disease and even being eradicated, so that there are few places left where such a relationship could be tested. Since there are also moral problems involved, I suspect that direct proof of the factors which are balancing the hemoglobin polymorphisms will never be forthcoming, except perhaps by means of some laboratory experiments which are technically not feasible at present. But making the assumption that ma-
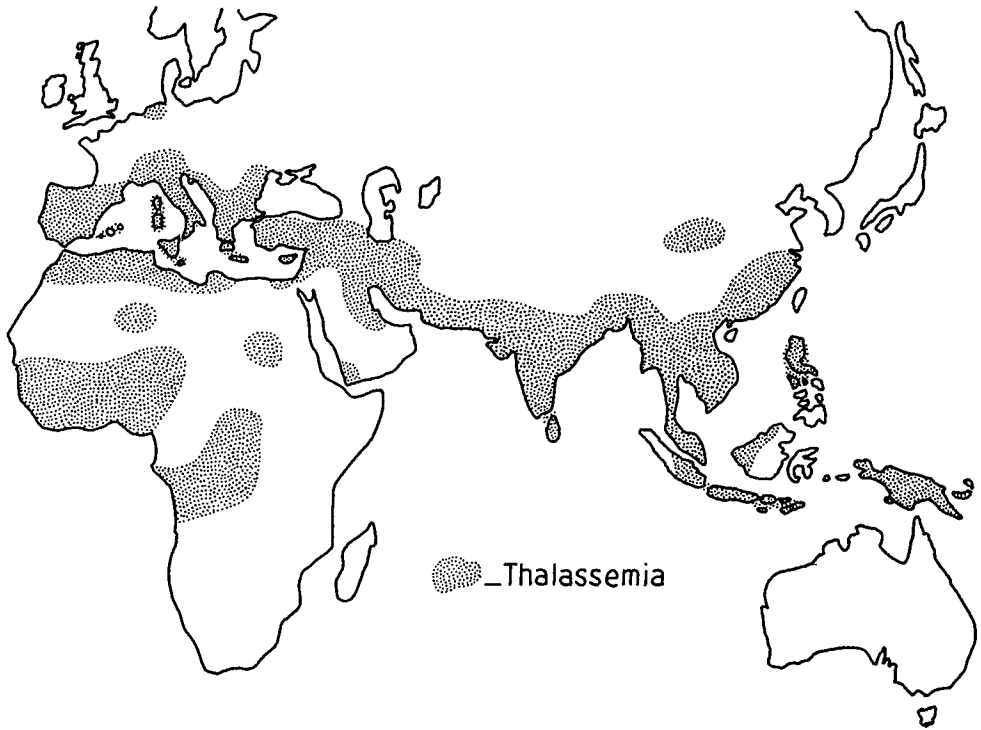
DISTRIBUTION OF THALASSEMIA

FIG. 1. The distribution of thalassemia in the Old World.

laria is the selective factor balancing these polymorphisms renders the entire distributions of these genes so intelligible that I think circumstantial evidence will prevail.

In the remainder of this paper malaria will be assumed to be the major factor which is balancing these polymorphisms. However, hemoglobin S seems to be balanced by falciparum malaria while this does not seem to be the case for the other abnormal hemoglobins (Livingstone, 1961). Thalassemia, for example, seems to convey some resistance to the other widespread species of malaria, quartan and benign tertian, which are due respectively to *Plasmodium malariae* and *P. vivax*. The fourth species of malaria parasite, *P. ovale,* is so rare that it need not be considered further.

Fig. 1 shows the distribution of thalassemia in the Old World. It only shows the presence of thalassemia and not frequencies since these data are quite incomplete. Thalassemia is the most widespread of the abnormal hemoglobins and its distribution in the Old World is almost as extensive as the distribution of endemic malaria, with the very striking exception of East Africa. A very few individuals who appear to be simultaneously heterozygous for hemoglobin S and a thalassemia gene have been discovered in East Africa. Strangely enough, on the edge of East Africa among the Arab and Greek communities of Khartoum, thalassemia apparently attains its highest frequencies in the world (Vella and Hassan, 1961; Vella and Ibrahim, 1961).

Occasionally, a thalassemia gene is encountered outside this area in Swiss, German, Japanese, English, and some others. On the map I have tried to indicate the areas where there appear to be a significant

number of cases or percentages of thalassemia. This association with few exceptions between endemic malaria and thalassemia, in addition to confirming the malaria hypothesis, indicates other factors in the population dynamics of the abnormal hemoglobins. Thalassemia, as a number of specific genes, has a higher mutation rate than the more specific abnormal hemoglobins; so that when malaria moves into an area, the first genes with an increased fitness which it is likely to encounter are thalassemia genes. Thus, as malaria has spread, it has brought thalassemia right behind it. For example, malaria appears to have spread rather recently through the islands of the southwest Pacific and is still spreading today through the Solomons. The Solomons and the New Hebrides are one of the few places in the tropics where there are anopheline mosquitoes and no malaria. Although malaria seems to have spread there from Southeast Asia, there has been little gene flow from the Indonesian populations to New Guinea. Hence in the New Guinea populations there are appreciable frequencies of thalassemia and another red cell gene, glucose-6-phosphate dehydrogenase deficiency, whose population genetics appears to be determined by approximately the same factors as thalassemia. In addition, there is a correlation between the amount of malaria and thalassemia in New Guinea (Curtain, Kidson, Gajdusek, and Gorman, 1962). Both are found in the lowlands but not in the highlands.

Turning now to the distribution of hemoglobin E, which is shown on fig. 2, it can be seen to have a rather restricted distribution on the mainland of Southeast Asia and in the Indonesian archipelago. This gene appears to have spread by gene flow out through Indonesia as far as Celebes and Timor. Since the mutation rate to this very specific chemical change would be much less than that to thalassemia, gene flow seems to be a reasonable assumption for this contiguous distribution. Among the Bugis of southern Celebes and some of the other peoples in this area hemoglobin O attains frequencies of about 2% and is almost unique to these populations (Lie-Injo and Sadono, 1958). As is the case with other hemoglobins with intermediate frequencies, hemoglobin O is found at the forefront of the wave of advance of one of the more widespread hemoglobins which are found in high frequencies, in this case hemoglobin E. Hemoglobin O, like other intermediate hemoglobins, appears to be a random mutant that in the absence of one of the hemoglobins with greater selective advantage such as hemoglobin E, has been selected for and increased in these peripheral populations.

In contrast to Africa and hemoglobin S, the distribution of hemoglobin E in Southeast Asia is more spotty and erratic. This is related to the fact that all species of human malaria are not as solidly endemic on every village and farm in Southeast Asia. This in turn is due primarily to the vectors of malaria there. On the mainland of Southeast Asia, *Anopheles minimus* is the major vector of malaria but in Malaya it is a similar species, *A. maculatus*. These vectors breed principally in small, cool, fast-flowing, sunlit streams and hence are found in great numbers only in the hilly regions of Southeast Asia, where the forest has been cut down. In contrast to Europe where malaria and marshes were intimately associated because of the nature of the anopheline vectors, in Southeast Asia the deltas and great river valleys are not particularly malarious. Thus, around Bangkok and in the delta of the Mekong River in South Viet-Nam there is little malaria and low frequencies of hemoglobin E. On the other hand, the foothills are intensely malarious and the more primitive slash and burn agriculturalists of these areas have high frequencies of hemoglobin E. No data on the frequency of hemoglobin E in the hunters of this area such as the Semang are available, but due to the ecology of the vectors, there is practically no malaria in the unbroken tropical forest. It is only when the forest has been cut down that *A. maculatus* becomes very frequent in Malaya and *A. minimus* elsewhere. *A. maculatus*, however, does not particularly prefer
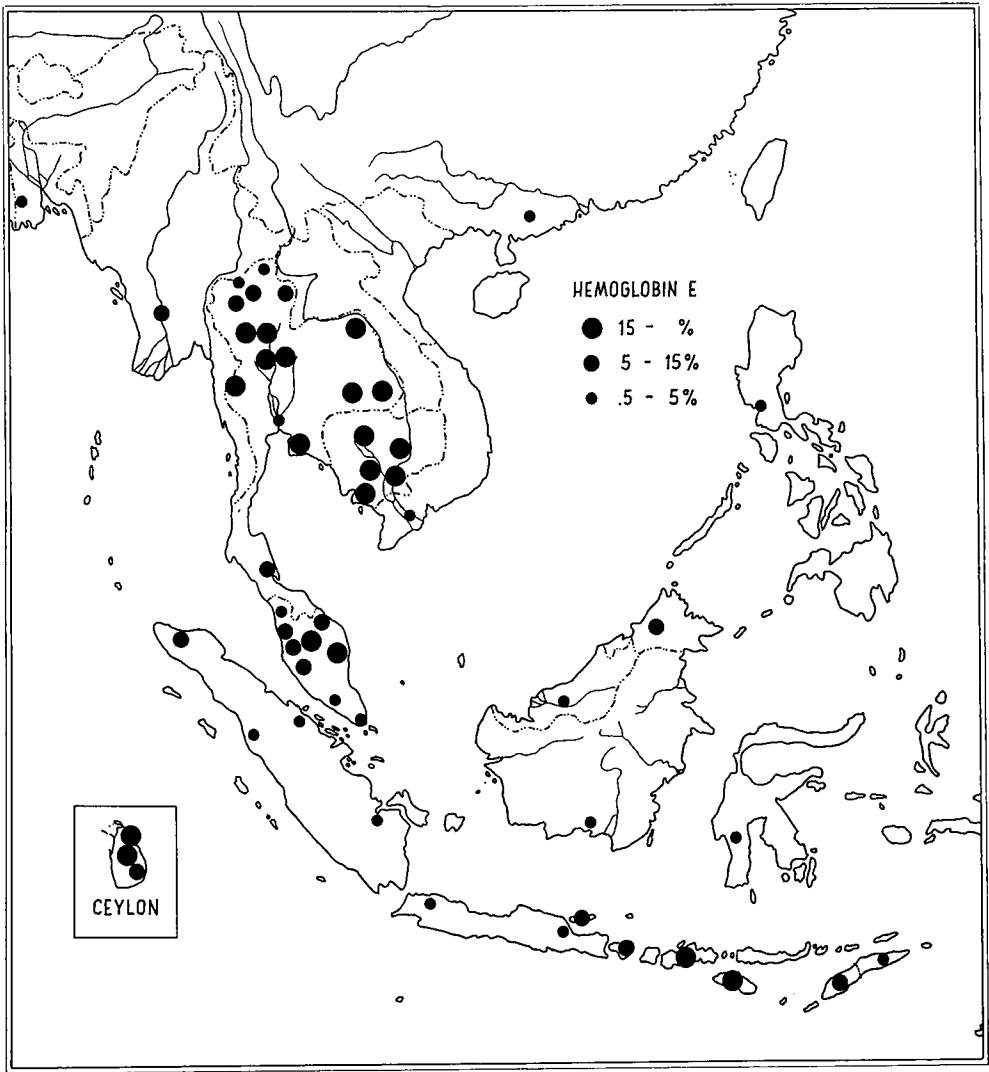
FIG. 2. Percentage incidence of individuals with hemoglobin E in the populations of Southeast Asia.[1]

human blood and can be diverted to biting cattle if these are available. Hence, in

primitive slash and burn agriculturalists with few cattle such as the Senoi there are high frequencies of hemoglobin E. But as more livestock are present and more human manipulation of the landscape occurs, malaria and $A.$ $maculatus$ tend to decrease.

In mainland Southeast Asia it has been recently demonstrated that $A.$ $balabacensis$ is an important vector in addition to $A.$ $minimus$ (Eyles, Wharton, Cheong, and Warren, 1964). Since this mosquito is associated with forests and appears most

numerous in villages surrounded by forest, it is perhaps responsible for much of the malaria in the sparsely populated areas. In the region of Cambodia where Eyles et al. (1964) worked, the parasite rates were very high and these populations also have very high frequencies of hemoglobin E.

In the swamps of parts of Malaya and into Indonesia, there are two other vectors of malaria, *Anopheles sundaicus* and *A. umbrosus*. *A. umbrosus* breeds in freshwater, shaded mangrove swamps and when the trees are cut down becomes scarce, while *A. sundaicus* is found only along the coast in brackish water. Thus, malaria is rather spottily distributed in Indonesia, but can become a very severe disease in places. On Borneo the efficient vectors of malaria are relatively absent and most of the malaria is due to *A. leucosphyrus* which breeds in shaded water and is associated with tropical forest. The Dyaks have less malaria and low frequencies of hemoglobin E, while the Muruts of North Borneo have more malaria and a higher frequency of hemoglobin E.

To the west hemoglobin E is occasionally encountered in Nepal and in Bengal there is a frequency of about 4%, but the greater part of India is characterized by high frequencies of hemoglobin S. Some hemoglobin E has been reported in Eti-Turks and Greeks, but it may be derived from other mutations and not represent gene flow from the east, although that possibility exists. With these exceptions there appears to be a strong border in the neighborhood of Calcutta with hemoglobin E predominating to the east and hemoglobin S to the west. This line coincides with the two general avenues of the penetration of agriculture into the South Asian tropical forest. Most of peninsular India and the Ganges River valley were populated by peoples from the west, while the Nagas and other tribes of the North East Frontier are linguistically and culturally related to peoples of Burma and southern China. However, it should be noted that few studies have been done on the Nagas or the other tribes of this area, but an assorted sample

of Burmese did have a high frequency of hemoglobin E.

The only groups to the west which have high frequencies of hemoglobin E are some Vedda villages on Ceylon. An earlier report (Aksoy, Bird, Lehmann, Mourant, Thein, and Wickremaisinghe, 1955) showed a very low frequency, but recently some villages are reported to have 30% hemoglobin E carriers (Wickremaisinghe, Ikin, Mourant, and Lehmann, 1963). The investigators imply that this indicates a genetic relationship to the Senoi but there is little evidence for it. Assuming that the genes are the same, it seems to indicate gene flow from Southeast Asia. Hemoglobin E is also found occasionally in the other inhabitants of Ceylon so this seems to be quite likely. It should be noted that the Veddas are now living in settled villages, have mixed considerably with the Tamils and Sinhalese, and the villages with the high frequencies of hemoglobin E are located in the north part of Ceylon which has hyperendemic malaria. Other studies of similar villages of Tamils and Sinhalese are needed, but since they are not considered to have "ethnological" significance they have not been investigated. To postulate that the presence of this gene in the Veddas and Senoi is the result of their common ancestry which goes back to a time when hunters and gatherers were sparsely distributed throughout Southeast Asia does not accord at all with our knowledge of the operation of natural selection on this gene.

Hemoglobin E then seems to have followed the spread of malaria through Southeast Asia. On the other hand, the spread of the conditions which gave rise to holoendemic malaria in India to the west of Calcutta appear to have come from Asia Minor. This part of India has hemoglobin D and thalassemia in the north, while hemoglobin S is found over most of the Indian peninsula, as is shown on fig. 3. Many gaps on the map are not necessarily areas where hemoglobin S is absent but just where no studies have been carried out. However, populations of the plains of Madras and Kerala have been examined
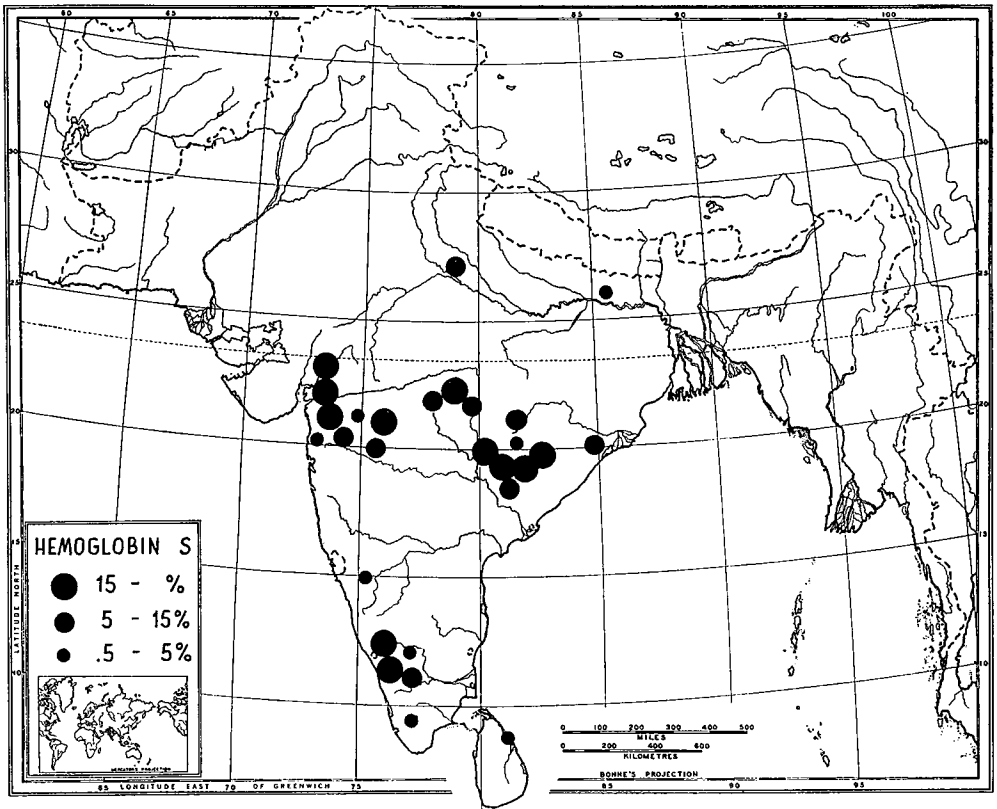
FIG. 3. Percentage incidence of individuals with hemoglobin S in the populations of India.

and no sickling is found. And strange as it may seem, these areas also have little malaria. The populations which have very high frequencies of hemoglobin S are found in the Western and Eastern ghats and the Bastar Hills. Here in these areas the vector of malaria, $A.$ $fluviatilis$, is like those of Southeast Asia in that it breeds in the little streams in the hills. These areas are characterized by holoendemic malaria while the coastal plains have little. On the Deccan Plateau of Mysore and Hyderabad malaria tends to occur in epidemics which are dependent on the breeding of the malaria vector, $A.$ $culicifacies$, which in turn is dependent on the monsoon. Although not many studies of abnormal hemoglobin have been done on the plateau, there seem to be lower frequencies.

The high frequencies of hemoglobin S

in South India also are found in tribal peoples, most of whom are slash and burn agriculturalists but some, like the Paniyans, verge on hunting and gathering. On the other hand, the castes, with the exception of the Mahars, seem to have much lower frequencies even in the same general areas, than the tribal peoples. This seems reasonable since malaria is not endemic in the cities of southern India because there is no "rainbarrel" vector there.

The question as to whether the hemoglobin S genes in India are derived from the same source as those in Africa is a real problem. Assuming that these are examples of the same mutation chemically speaking, then the restricted distribution of hemoglobin S in Africa, the Middle East, and India seems to indicate that this gene is a rather rare mutant. None of the popula-

HEMOGLOBIN C

● 15 - %

● 5 - 15%

• .5 - 5%
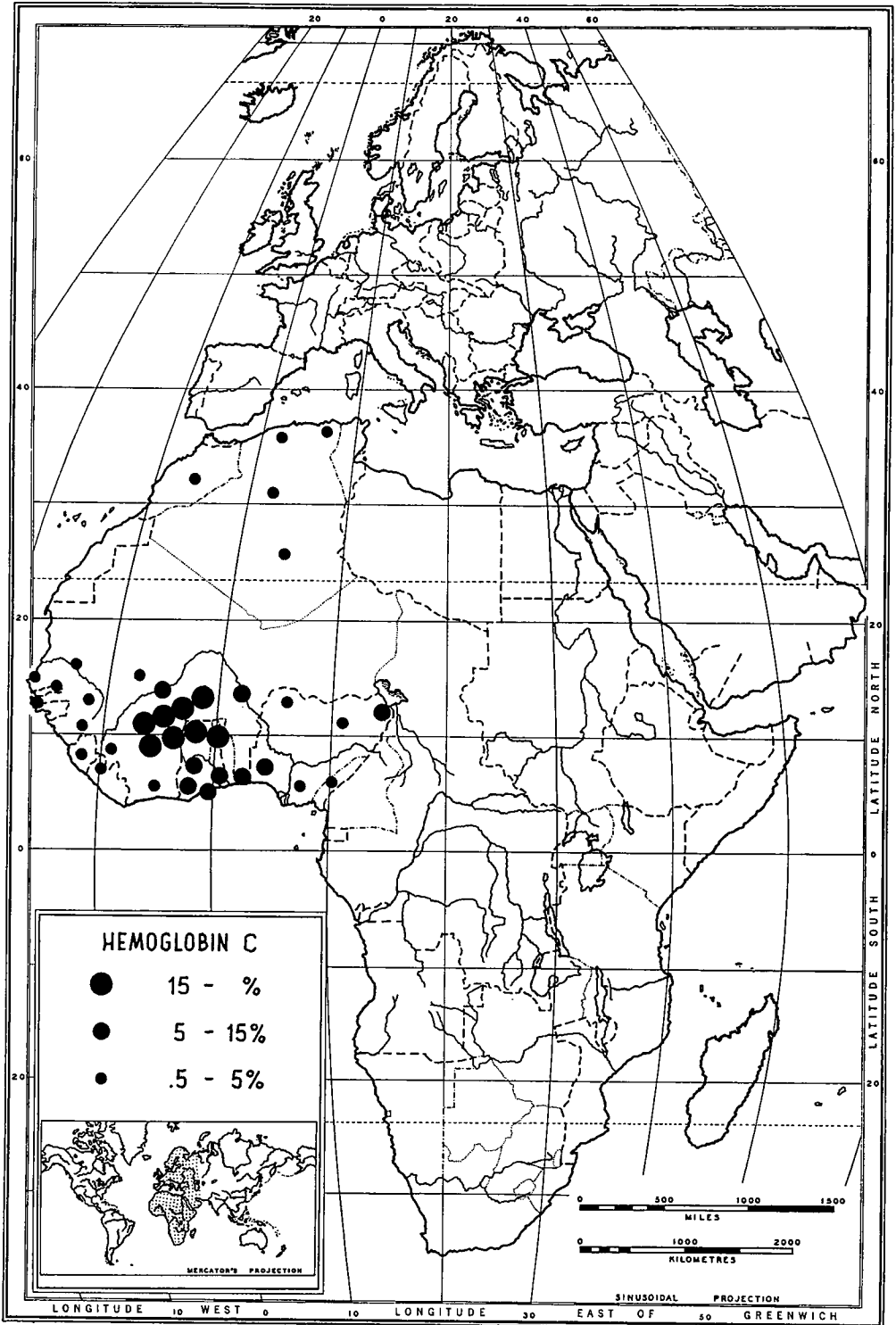
tions with high frequency are very distant from others and hence gene flow has presumably occurred.

Turning to Africa and the Middle East, fig. 4 shows the distribution of hemoglobin C. This gene is restricted to West and North Africa and the high frequencies in West Africa are found in the more primitive tribes that speak Gur languages. The one exception is the Wakura who inhabit the northernmost mountains in the Cameroons. They appear to be long resident in the area and are surrounded by the Kanuri and Shuwa Arabs among whom there are extremely high frequencies of hemoglobin S and no hemoglobin C. One of the problems associated with hemoglobin C is whether it is replacing hemoglobin S or vice versa, or perhaps there is an equilibrium frequency with both present. The fitness values of the genotypes involved in this triallelic system undoubtedly vary throughout this area, but I think a good case can be made for the hemoglobin S gene replacing hemoglobin C. I think it can be shown that hemoglobin S will replace almost any other abnormal hemoglobin gene or thalassemia.

From the fact that the frequency of hemoglobin S carriers is 30 to 40% in many populations in this part of the world, one can estimate that the fitness of the AS heterozygote must be 1.2 to 1.3 times the fitness of the normal homozygote, assuming that the sickle cell homozygote has a fitness of zero. On the other hand, since thalassemia only occurs up to 25% of some populations, the fitness of the thalassemia heterozygote must be about 1.1 times that of normals. Homozygotes for hemoglobin C do not have the very serious anemia of either sickle cell or thalassemia homozygotes. If we estimate the fitness of the homozygous C genotype at from 0.6 to 0.8 that of the normal genotype, then, since a C gene frequency of 0.15 is as high as the equilibrium C gene frequency could possibly be with only C present, the maximum fitness value of the hemoglobin C heterozygote would be about 1.08 times the normal fitness, which we arbitrarily set as 1.00. Given these fitness values, we can attempt to determine which of these genes, hemoglobin C, S, or thalassemia, will replace the others or whether instead there is an equilibrium point with more than one of them present. The general conditions for the existence of an equilibrium with three genes present have been solved by Li (1955), Kimura (1956), Penrose, Smith, and Sprott (1956), and Mandel (1959), and Bodmer and Parsons (1961) have solved the conditions as to whether any one of a series of alleles will increase when introduced into a population at a low frequency. However, the differential equations involving gene frequency change at a multiallelic locus are non-linear and non-homogeneous and hence a general solution is not possible by the methods of classical mathematical analysis; but one can use a computer to obtain specific solutions to these equations. I was rather amazed to find that the sickle cell gene will increase under a great variety of conditions. It will completely replace the hemoglobin C gene and will take over from the thalassemia gene although there is a stable equilibrium with about 15% sickle cell genes and 2% thalassemia. In both cases it takes about 50 generations or about 1,000 years for the sickle cell gene to replace the C or thalassemia gene. In evolutionary perspective this is a very short time span; in fact, it is about the most rapid rate at which one gene can replace another. Thus, what seem to be the most plausible estimates of the fitness values of the abnormal hemoglobin genotypes imply that the sickle cell gene is extremely predatory and will eliminate other such genes.

This analysis has many implications for the distributions of the S and C genes. In West Africa their distributions have raised

---

Fig. 4. Percentage incidence of individuals with hemoglobin C in the populations of North and and West Africa.

HEMOGLOBIN S

● 15 - %

● 5 - 15%

• .5 - 5%
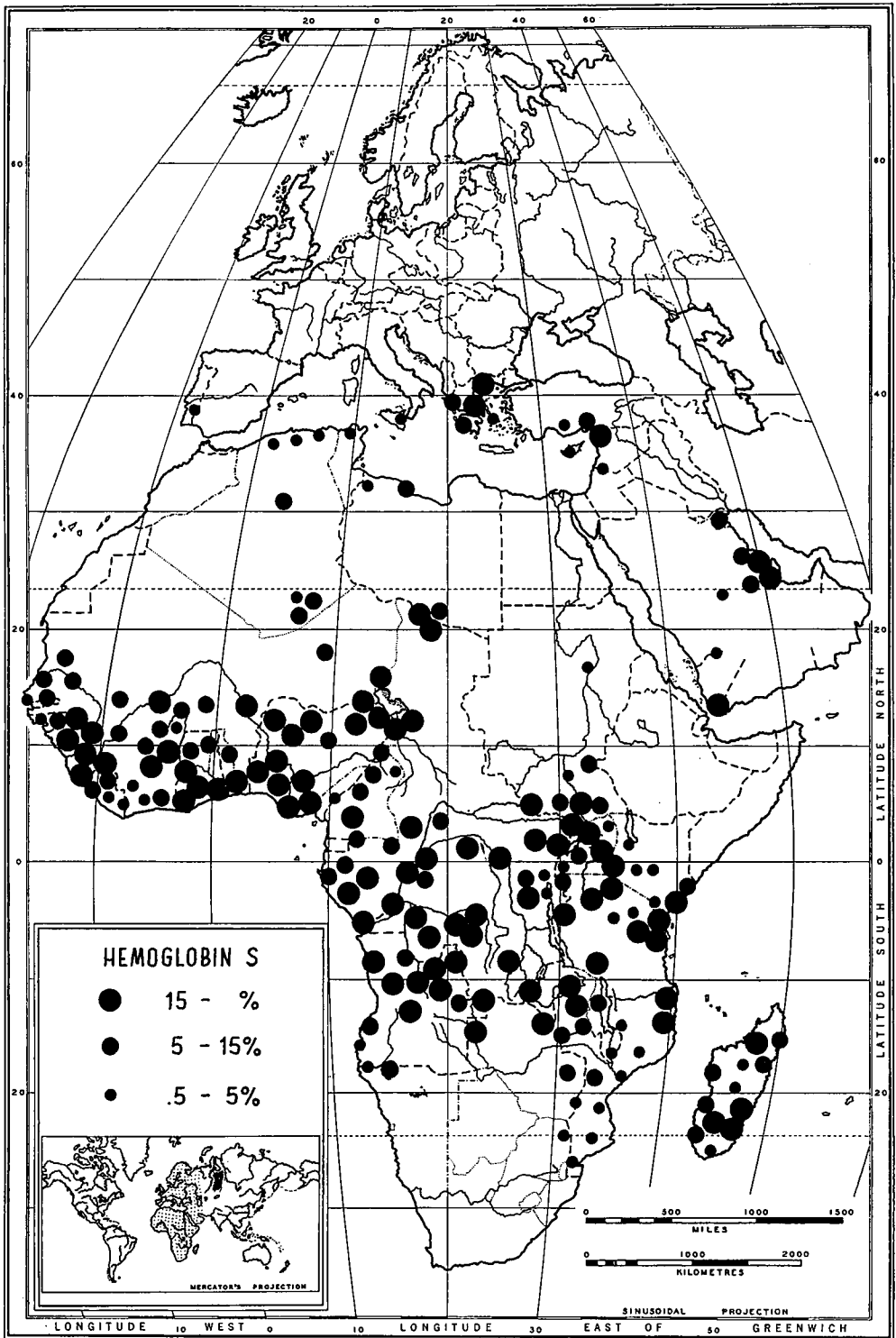
the problem as to whether there is an equilibrium with both genes present as Allison (1956) has maintained, or whether one gene is replacing the other. According to this analysis the S gene should be replacing the C gene, and this seems reasonable. Since there are populations in the western part of West Africa with low frequencies of both genes, it appears that the S gene was introduced into this area from the east and north, and historically it is known that a great deal of gene flow did occur in this direction. Since the highest frequencies of the C gene are found in the rather primitive Gur-speaking peoples who have received less such gene flow, it appears that the C gene was present here before the S gene and is now being replaced.

Fig. 5 shows how the hemoglobin S gene has spread throughout the Old World. In East Africa the S gene has replaced all the other abnormal hemoglobins, while to the north in the Mediterranean area, S is only found in high frequencies in North Africa and parts of Greece and Turkey. However, Barnicot, Allison, Blumberg, Deliyannis, Krimbas, and Ballas (1963) have shown that there is an inverse correlation between the S and thalassemia gene frequencies in the malarious areas of Greece, which indicates that they are in competition and that one is replacing the other. The S gene does not seem to have reached Sardinia which seems expectable since it is one of the more isolated areas of Italy with less contact with and presumable gene flow from the outside.

The fact that the S gene will increase rapidly in a malarious environment no matter what other genes are present implies that any mutant S gene would most likely increase. Thus, the fact that there are many malarious areas without high frequencies of the S gene and that the high frequencies of this gene are in rather contiguous populations indicates that the S gene is a rare mutant and that the distribution of this gene is due primarily to gene flow among the populations within which it occurs.

CONCLUSIONS AND SUMMARY

The preceding interpretation of the distributions of the abnormal hemoglobins in terms of the factors which we know control gene frequency change has several general implications. First, natural selection is the major factor determining differences in gene frequency, and for the abnormal hemoglobins this natural selection is determined by the presence of holoendemic malaria and its high cost of life. The evolution of this type of malaria is in turn dependent upon a sedentary way of life for which agriculture seems to be a necessary prerequisite in most environments. Agriculture and hence high frequencies of abnormal hemoglobins are relatively recent events in human evolution. The penetration of agriculture into the tropical regions of Asia came from two centers, China and the Middle East. The penetration from China through Southeast Asia to Indonesia is associated with the spread of hemoglobin E, while the spread of agriculture from the Middle East through peninsular India and Africa is associated with the spread of the hemoglobin S gene. Where these two movements met in the vicinity of Calcutta is also the border between the high frequencies of these genes, with S on the west and E to the east.

We have learned a great deal about the evolutionary process from the abnormal hemoglobins since they are the first genetic system in man to which we can apply genetic theory in any detailed way. They also emphasize the important relationship between man's cultural and biological evolution. The evolution of gene frequency differences has often been called "race formation." With reference to abnormal hemoglobin, such a label seems not only grossly inappropriate but false when the

←

FIG. 5. Percentage incidence of individuals with hemoglobin S in the populations of Africa, Europe, and Southwest Asia.

implications of the concept of race are taken into consideration. In addition, the distributions of the abnormal hemoglobin genes are not at all related to the traditional races of man, although one can still read that the sickle cell gene is a Negroid character. With reference to the hemoglobins, Coon (1962: 663) has recently stated, "To me, at least, it is encouraging to know that biochemistry divides us into the same subspecies that we have long recognized on the basis of other criteria," a statement which I think the detailed arguments of this paper have shown to be false with respect to the hemoglobins. If natural selection is considered to be one of the major factors contributing to the gene frequency differences which exist among the hundreds of thousands of human breeding isolates, then the gene frequencies will vary with the intensity of selection as do the abnormal hemoglobins and not with "race." If, as Mayr (1963) has recently emphasized, subspecies are artificial units of the classifier and not major units of evolution, it is time anthropologists stopped considering them as such. Between the "natural" units of species and breeding population, there is no natural unit according to the genetic theory, and to attempt to reconcile race with the genetic theory of evolution is to use old concepts to express new ideas which leads to nothing but confusion since we all learn what race basically is in childhood. In this paper I have attempted to describe and explain the distribution of one particular genetic system in man. This has been done without using the concept of race; I think the same can be done for any genetic system.

LITERATURE CITED

AKSOY, M., G. W. G. BIRD, H. LEHMANN, A. E. MOURANT, H. THEIN, AND R. L. WICKREMAI-SINGHE. 1955. Hemoglobin E in Asia. J. Physiol., 130: 56P–57P.

ALLISON, A. C. 1954a. Protection afforded by sickle-cell trait against subtertian malarial infection. British Med. J., 1: 290–294.

——. 1954b. The distribution of the sickle-cell trait in East Africa and elsewhere, and its apparent relationship to the incidence of sub-

tertian malaria. Trans. Roy. Soc. Trop. Med. and Hygiene, 48: 312–318.

——. 1954c. Notes on sickle-cell polymorphism. Ann. Human Gen., 19: 39–57.

——. 1956. Population genetics of abnormal human hemoglobins. Acta Gen. et Statistica Med., 6: 430–434.

BARNICOT, N. A., A. C. ALLISON, B. S. BLUMBERG, G. DELIYANNIS, C. KRIMBAS, AND A. BALLAS. 1963. Hemoglobin types in Greek populations. Ann. Human Gen., 26: 229–236.

BODMER, W. F., AND P. A. PARSONS. 1961. The initial progress of new genes with various genetic systems. Heredity, 15: 283–299.

CEPPELLINI, R. 1957. I meccanismi evolutivi nelle popolazioni umane. Suppl. a La Ricerca Scientifica, 27: 3–23.

CHOREMIS, C., P. FESSAS, C. KATTAMIS, STA-MATOYANNOPOULOS, L. ZANNOS-MARIOLEA, A. KARAKLIS, AND G. BELIOS. 1963. Three inherited red-cell abnormalities in a district of Greece, thalassemia, sickling, and glucose-6-phosphate-dehydrogenase deficiency. Lancet, 1: 907–909.

COON, C. S. 1962. The origin of races. Knopf, New York.

CURTAIN, C. C., C. KIDSON, D. C. GAJDUSEK, J. G. GORMAN. 1962. Distribution pattern, population genetics and anthropological significance of thalassemia and abnormal hemoglobins in Melanesia. Amer. J. Phys. Anthrop., 20: 475–483.

DELBROUCK, J. 1958. Contribution a la genetiquede la sicklemie. Ann. Soc. Belge de Med. Trop., 38: 103–133.

EYLES, D. E., R. H. WHARTON, W. H. CHEONG, AND McW. WARREN. 1964. Studies on malaria and Anopheles balabacensis in Cambodia. Bull. World Health Org., 30: 7–21.

FIRSCHEIN, L. 1961. Population dynamics of the sickle-cell trait in the Black Caribs of British Honduras, Central America. Amer. J. Human Gen., 13: 233–254.

INGRAM, V. M. 1963. The hemoglobins in genetics and evolution. Columbia University Press, New York.

KIMURA, M. 1956. Rules for testing stability of a selective polymorphism. Proc. Nat. Acad. Sci., 42: 336–340.

LI, C. C. 1955. The stability of an equilibrium and the average fitness of a population. Amer. Nat., 89: 281–296.

LAMBOTT-LEGRAND, J., AND C. LAMBOTTE-LEGRAND. 1958. Notes complementaire sur la drepanocytose. Ann. Soc. Belge de Med. Trop., 38: 45–54. I. Sicklemie et malaria.

LIE-INJO, L. E., AND M. D. SADONO. 1958. Haemoglobin O (Buginese X) in Sulawesi. Brit. Med. J., 1: 1461–1463.

LIVINGSTONE, F. B. 1961. Balancing the human hemoglobin polymorphisms. Human Biol., **33**: 205–219.

MANDEL, S. P. H. 1959. The stability of a multiple allelic system. Heredity, **13**: 289–302.

MAYR, E. 1963. Animal species and evolution. Harvard Univ. Press, Cambridge.

NEEL, J. V. 1951. The population genetics of two inherited blood dyscrasiias in man. Cold Spring Harbor Symp. Quant. Biol., **15**: 141–158.

PENROSE, L. S., S. M. SMITH, AND D. A. SPROTT. 1956. On the stability of allelic systems, with special reference in haemoglobins A, S, and C. Ann. Human Gen. **21**: 90–93.

RAPER, A. B. 1956. Sickling in relation to morbidity from malaria and other diseases. Brit. Med. J., **1**: 965–966.

SHIBATA, S. 1961. Hemoglobinopathy, with special reference to the abnormal hemoglobins found in Japan. Bull. Yamaguchi Med. School, **8**: 197–207.

THOMPSON, G. R. 1962. Significance of haemoglobins S and C in Ghana. British Med. Jour., **1**: 682–685.

VELLA, V., AND M. M. HASSAN. 1961. Thalassemia major in a Sudanese Arab family. J. Trop. Med. and Hygiene, **64**: 199–201.

—— AND S. A. IBRAHIM. 1961. The frequency of thalassemia minor in a Greek community. J. Trop. Med. and Hygiene, **64**: 202–206.

WICKREMAISINGHE, R. L., E. W. IKIN, A. E. MOURANT, AND H. LEHMANN. 1963. Blood groups and haemoglobins of the Veddas. J. Roy. Anthrop. Inst., **93**: 117–125.