Bull. Org. mond. Santé }
Bull. Wld Hlth Org. } 1966, 34, 379-393

Population Genetics of Haemoglobin Variants, Thalassaemia and Glucose-6-Phosphate Dehydrogenase Deficiency, with Particular Reference to the Malaria Hypothesis

M. SINISCALCO,[1] L. BERNINI,[1] G. FILIPPI,[2] B. LATTE,[3] P. MEERA KHAN,[1] S. PIOMELLI[4] & M. RATTAZZI[1]

The authors report data on the genetic distribution of thalassaemia and of glucose-6-phosphate dehydrogenase deficiency in the populations of certain Sardinian villages, many of which are not only of great antiquity but have maintained isolation for very long periods and therefore possess the following three requirements for suitability for investigation of the possible interrelationships among malaria, thalassaemia and G-6-PD deficiency: a reasonable degree of ethnic homogeneity, availability of reliable demographic data, and availability of malaria-free populations of adequate size and of ethnic background and genetic isolation similar to those of the malarial populations.

Investigations including more than 6000 observations in 52 villages demonstrated a positive correlation between the incidences of thalassaemia and G-6-PD deficiency. It is suggested that the genotype that carries thalassaemia and/or the enzyme deficiency may have a high adaptive value in a malarial environment.

It is concluded that there is a need further to investigate human genetic structure and the biological fitness of the principal genotype combinations in both existing environments and those that will result from continued cultural evolution.

INTRODUCTION

The present status of knowledge of haemoglobin variants, thalassaemia and glucose-6-phosphate dehydrogenase (G-6-PD) deficiency has been summarized on several occasions in recent years. Excellent reviews of all aspects of these questions are now available (Allison, 1965; Baglioni, 1962; Fessas,1965; Ingram, 1963; Itano, 1965; Motulsky, 1965; Rucknagel, 1964; Silvestroni & Bianco, 1963).

Consideration of the recent exhaustive accounts of Silvestroni & Bianco (1963) on the world distribution of haemoglobin variants and of thalassaemia, that of Motulsky (1965) on G-6-PD deficiency, together with the elegant monograph of Rucknagel & Neel (1961) on the dynamics, at a population level, of the genes controlling these conditions, and the discussions of the same subject by Livingstone (1964) and by Allison (1965) may serve as ideal introductions to the subject of the present report—that is, the hypothesis that malaria may have been the common ecological factor that was responsible for the selection of these three groups of inherited abnormalities of the red cells.

The idea that all individuals might not be equally liable to malarial infection was first proposed by Haldane (1949) to explain the preponderance of thalassaemia in the Mediterranean basin. A few years later, Allison (1954) reported the interesting finding that persons who carry the sickle-cell trait are indeed more resistant to subtertian malaria than those who do not. Although this claim has been disputed in some instances, it is now supported by a most impressive body of evidence, including data of the three following kinds (Allison, 1965): (1) direct demonstration, in areas where malignant malaria is still endemic, that young children who are heterozygous for the sickle-cell gene have lower *Plas-*

[1] Department of Human Genetics, University of Leiden, Leiden, Netherlands.
[2] Department of Genetics, University of Rome, Rome, Italy.
[3] Town Hospital, Nuoro, Italy.
[4] Department of Genetics, University of Rome. Present address: Department of Pediatrics, New York University–Bellevue Medical Center, New York, N.Y., USA.

modium falciparum parasite counts than do children without this trait; (2) the finding of a low incidence of sickling carriers among the cases of fatal malarial infections observed in the same areas; and (3) the overlap between the world distribution of the sickle-cell trait and that of malignant malaria.

Motulsky (1960) was the first to report population data that showed that the distribution of G-6-PD deficiency in the Eastern Hemisphere also overlaps that of malignant malaria, and suggested that even this well-known example of sex-linked polymorphism owes its establishment to the higher fitness of enzyme-deficient genotypes in a malarial environment.

On the other hand, studies of the malaria parasite counts in groups of normal and enzyme-deficient children yielded contradictory results. Thus, while Allison & Clyde (1961), in East Africa, and Harris & Gilles (1961) in West Africa found significantly lower *P. falciparum* counts in young enzyme-deficient children, Kruatzachue and his co-workers (1962), Motulsky (1965), and Edington & Watson-Williams (1965) failed to do so.

Since, however, it is known that, for instance, a protein-deficient diet (Pérez et al., 1964) and hypothyroidism (personal unpublished data) can lower G-6-PD activity, it is uncertain how much of this disagreement may be due to misclassification of the G-6-PD phenotypes, which reasonably may be expected, especially when dealing with indigenous African populations, in which enzyme deficiency is not as complete as in Caucasians, and when only screening tests have been used for the diagnosis.

Equally contradictory have been the conclusions of studies on malarial parasite counts in the heterozygous carriers of haemoglobin C and haemoglobin E genes (Edington & Laing, 1957; Thompson, 1962; Edington & Watson-Williams, 1965; Brumpt & Brumpt, 1958; Kruatzachue et al., 1961). Investigations of these types have been impossible to perform on other haemoglobin variants because of their rarity, or on the different forms of thalassaemia because their correct diagnosis, which requires elaborate laboratory studies, is unreliable under field-work conditions. Moreover, the common occurrence, in the primitive areas where malaria is still prevalent, of environmental and biological stress factors may distort the haematological picture of thalassaemia.

Thus, it is not surprising that the only evidence for the relationship between malaria, thalassaemia and G-6-PD deficiency should come exclusively from population studies showing a positive correlation between the incidence of these genes and malaria morbidity or, rather, past malaria morbidity; since, for the reasons outlined above, the correct classification of these abnormalities of the red cells is difficult in primitive areas where malaria is still endemic. Nevertheless, even these types of studies present difficulties, since they can have very little value unless the following three essential requirements are met: (1) a reasonable degree of ethnic homogeneity among the " genetic isolates " chosen for the study; (2) the availability of accurate historical data as well as general vital statistics, data on the malaria morbidity, mating patterns, and rates of consanguinity in order to be able to estimate the relative importance of drift, migration and natural selection as potential causes of any genetic heterogeneity that might be found among neighbouring isolates; and (3) the availability of control populations within the malarial areas under consideration; that is, the existence of " malaria-free islands " or human settlements of appreciable size, inhabited by persons of the same ethnic group, who have been living in genetic isolation for a very long time. Populations of nomadic habits are clearly useless for investigations of this type.

Since these conditions are met only very rarely, it is impossible to attempt an over-all evaluation of the "malaria hypothesis" from the data available on the world distribution of the genes under discussion here. Instead, we prefer to report here, for the first time in full, the data that we collected in Sardinia and have so far published only in part (Siniscalco et al., 1961; Siniscalco, 1964; Adinolfi et al., 1960), since we feel that they represent one of the few sets of data among the population studies so far published on the subject that meet the requirements outlined above.

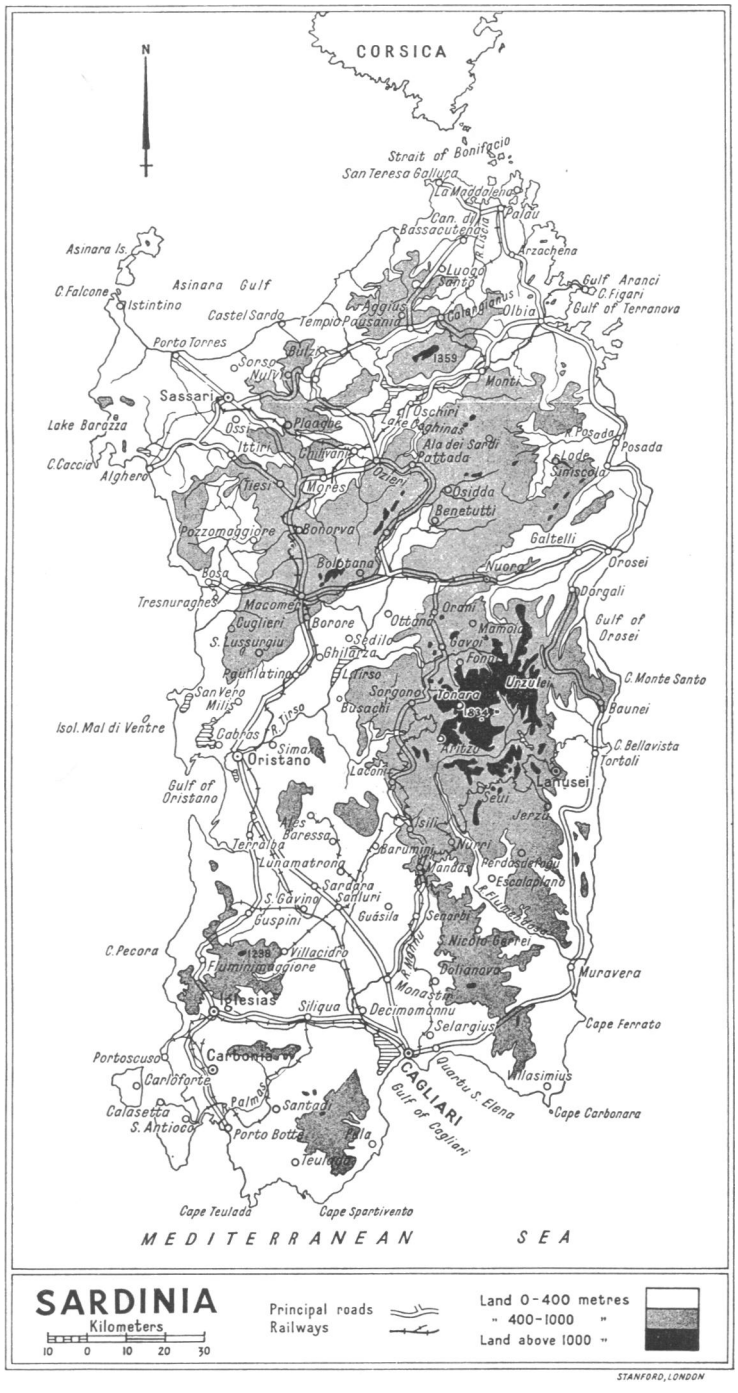
HISTORICAL BACKGROUND

The island of Sardinia is the ideal place for such population studies. The degree of isolation to be found there is still very high for most of its villages and, as shown in Fig. 1, its geography is such that, within a few hundred square miles, one can easily find isolated settlements with very high malaria morbidity in the past and others that have been practically always free from this disease.

Furthermore, it can reasonably be assumed that the populations of many Sardinian villages must have remained free from external admixtures over a very long period. Ancient Greek colonization was, in fact, limited to Olbia (the northern portion of the island), while the Romans and Carthaginians only exploited

FIG. 1

LOCATIONS AND ALTITUDES ABOVE SEA-LEVEL OF THE VILLAGES IN SARDINIA CITED IN THE TEXT [a]



[a] Reproduced from Logan, J. A. (1953) *The Sardinian project*, with the permission of the Johns Hopkins Press, Baltimore, Md., USA.

the coastal regions for grain, as is proved by the very localized areas along the south-west coast in which archaeological remains of their towns can now be found, as at the excavations at Nora. Later, the Vandals and Goths simply overran the island, which subsequently became the scene of struggles among the Pisans, Genovese and Saracens, none of whom, however, were there in sufficient numbers or for a long enough time to alter, significantly, the genetic structure of the autochthonous population. Even the Spanish, who ruled the island from 1297, when Pope Boniface VIII awarded it to James II of Aragon, never really cared to penetrate the rocky paths leading to the interior of the island, which remained half-forgotten, with its primitive villages consisting of huts clustered around the ancient nuraghi, and where little social or economic change took place until the accession of the House of Savoy in 1720.

The nuraghi, those beautiful and impressive Bronze Age stone constructions, are scattered in thousands all over the island, showing how well organized the Sardinians were, even in prehistoric times, and how concerned with defence against invaders. It can hardly be denied, however, that as a result of these invasions, a few sets of " external " genes must have entered the Sardinian genetic pool from time to time.

Malaria has probably been endemic in Sardinia since prehistoric times, according to the Roman historian Livy, and has remained one of the most important causes of infant mortality until the beginning of the present century. This disease was not eradicated completely until after the Second World War, when the Rockefeller Foundation, by a massive and well-planned anti-malaria campaign (Logan, 1953), successfully completed the pioneering work of Fermi and Missiroli, the distinguished Italian malariologists who had struggled against the disease in Sardinia for decades and made available to posterity the most accurate and complete information that could be desired on the malaria morbidity of every Sardinian village. This information, together with the very ancient church records and the first-class vital statistics on the island that have been available since the Savoy accession in 1720, enabled us to draw valid conclusions about the population structure and dynamics of the Sardinian isolates chosen for our studies.

When we began our investigations, the existence of thalassaemia and G-6-PD deficiency in Sardinia already had been well established (Carcassi, Ceppellini & Pitzus, 1957; Larizza et al., 1958), and the preliminary studies of Ceppellini (1955) had shown that the incidence of thalassaemia in two non malarial villages in the Gennargentu Mountains was strikingly low as compared with that found in two lowland villages in formerly very malarial parts of the eastern coast of the island.

We repeated these studies in a total of 19 villages and extended them to G-6-PD deficiency (Siniscalco et al., 1961) and were able to demonstrate that there was, indeed, a very close positive correlation between the present-day frequencies of thalassaemia and G-6-PD deficiency and former malaria morbidity as reported by Fermi (1938).

Additional studies of the same kind have been made in the past few years; new data are now available and are discussed in detail below.

SUMMARY OF THE SARDINIAN POPULATION DATA

Table 1 shows the frequency of the gene for G-6-PD deficiency, Gd(−), and that for β-thalassaemia, Th(+), in 52 Sardinian villages, based on more than 6000 observations of unrelated individuals. The estimated frequencies of the two traits refer to the youngest generation, since the data for each village were obtained from a random sample of its schoolboys.

These estimates are better summarized in Fig. 2, in which a positive correlation between the frequency of thalassaemia and that of G-6-PD deficiency is evident. This correlation seems to be described reasonably well by a straight regression line up to a maximum frequency level of about 24%. The correlation fades off above this level, since the gene for thalassaemia, being lethal in the homozygous condition, can never reach equilibrium levels as high as those that are possible for the much less unfavourable gene for the enzyme deficiency.

Basic considerations of population genetics (Livingstone, 1964) make it evident that:

(1) Levels of gene frequency such as those reported for the villages in the Sardinian plains can be explained only by assuming a higher fitness of the heterozygous genotype for thalassaemia and of the heterozygous and, perhaps, the hemizygous and homozygous genotypes for enzyme deficiency.

(2) A " migration " hypothesis is grossly inadequate to account for the very high frequencies of these traits found in the plains, since they could not be expected, even if one were to assume, against all historical evidence, that the autochthonous populations had been totally replaced by equally numerous

TABLE 1

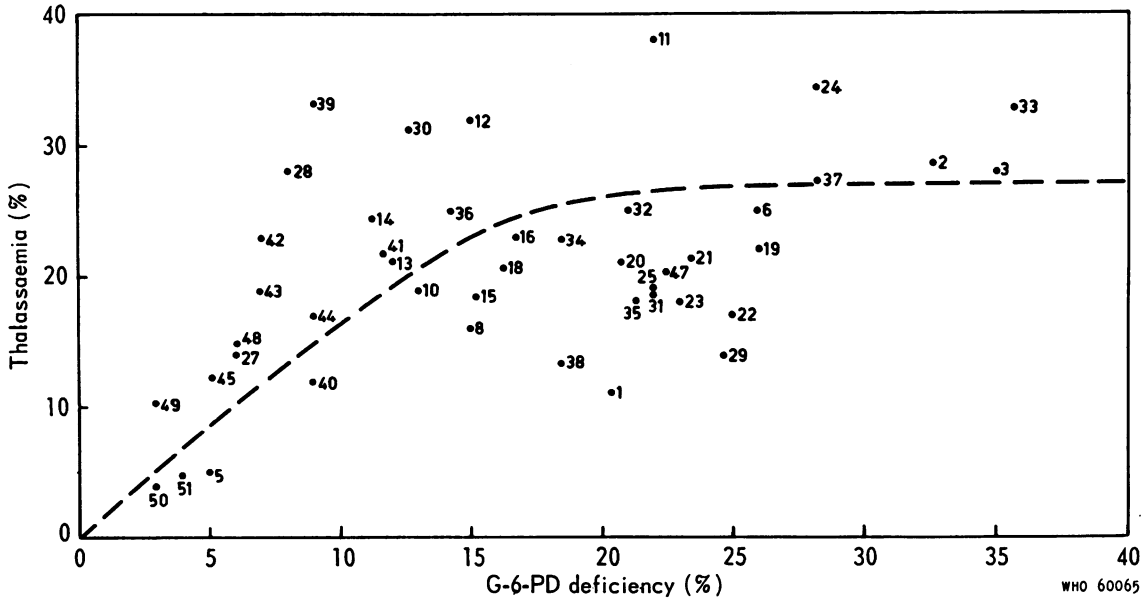
INCIDENCE OF GLUCOSE-6-PHOSPHATE DEHYDROGENASE DEFICIENCY AND THE β-THALASSAEMIA TRAIT IN 52 SARDINIAN VILLAGES

Village [a]	Altitude (metres)	Gd(—) [b]		TH(+) [c]	
		Number tested	Percentage positive	Number tested	Percentage positive
1. Assemini	6	108	20.4	108	11.0
2. Marrubiu	7	98	32.6	98	28.6
3. Cabras	9	200	35.0	100	28.0
4. Terralba	9	100	30.0	—	—
5. Carloforte	10	99	5.0	99	5.0
6. Decimomannu	10	100	26.0	100	25.0
7. S. Giusta	10	42	30.9	—	—
8. Pula	15	100	15.0	100	16.0
9. Tortoli	15	50	16.0	—	—
10. Orosei	19	180	13.0	308	18.8
11. Torpe	24	100	22.0	100	38.0
12. Irgoli	26	100	15.0	100	32.0
13. Galtelli	40	175	12.0	235	21.2
14. Siniscola	42	195	11.3	97	24.4
15. Barisardo	50	98	15.3	98	18.4
16. Teulada	50	101	16.9	100	25.0
17. S. Gavino	53	100	26.0	—	—
18. Capoterra	54	92	16.3	92	20.6
19. Siliqua	66	100	26.0	100	22.0
20. Vallermosa	70	86	20.9	86	21.0
21. Monastir	83	94	23.4	94	21.2
22. Nuraminis	86	100	25.0	100	17.0
23. Villamar	108	100	23.0	100	18.0
24. Guspini	137	99	28.2	99	34.4
25. Domusnovas	152	100	22.0	100	19.0
26. Gonnosfanadiga	156	49	24.5	—	—
27. Usini	190	99	6.1	99	14.2
28. Ottana	195	72	8.0	72	28.0
29. Senorbi	204	101	24.7	101	13.8
30. Tresnuraghes	257	86	12.7	86	31.4
31. Sedilo	288	100	22.0	96	18.8
32. Serrenti	307	100	21.0	100	25.0
33. Arbus	311	95	35.7	95	32.8
34. Abbasanta	315	97	18.5	92	22.8
35. Dualchi	321	75	21.3	100	18.0
36. Suni	333	98	14.3	100	25.0
37. Lode	335	820	28.2	820	27.6
38. Gergei	374	92	18.5	92	13.2
39. Borore	399	100	9.0	99	33.2
40. Benetutti	406	100	9.0	100	12.0
41. Bolotana	472	93	11.8	93	21.4
42. Luras	508	100	7.0	98	23.0
43. Lula	521	100	7.0	100	19.0
44. Isili	523	100	9.0	100	17.0
45. Bitti	549	193	5.1	193	12.2
46. Lanusei	595	100	4.0	—	—
47. Ala dei Sardi	663	80	22.5	80	20.0
48. Orune	745	97	6.1	97	14.4
49. Gavoi	777	98	3.0	98	10.2
50. Desulo	891	313	3.0	320	3.8
51. Tonara	935	148	4.0	102	4.8
52. Fonni	1 000	100	3.0	—	—

[a] These villages are arranged and numbered in order of their increasing altitude. The data were obtained from random samples of the schoolboys of each village.

[b] Glucose-6-phosphate dehydrogenase deficiency. [c] Thalassaemia.

FIG. 2
DISTRIBUTION OF G-6-PD DEFICIENCY AND OF THE THALASSAEMIA TRAIT IN SARDINIA [a]



[a] Estimates of the gene frequencies obtained from random samples of schoolboys. The numbers identify the villages listed in Table 1.

groups of immigrants, all of whom were carriers of thalassaemia and/or G-6-PD deficiency.

(3) If the genetic heterogeneity between the lowland villages and those in the high mountains is attributable to different adaptive values of the carrier genotypes, malignant malaria is the obvious ultimate factor, since the two environments are known to have differed from each other for centuries and, until only about twenty years ago, almost exclusively in respect to mortality and morbidity from malaria.

The negative correlation with altitude above the level of 400 m is clearly demonstrated in Fig. 3, where the average frequencies (± 3 times their sampling errors) of the two traits are reported for each group of villages of similar altitude. This correlation is therefore positive when gene frequencies are compared with the relative incidence of malaria, as we demonstrated in a series of villages for which direct estimates of past malaria morbidity were available (Siniscalco et al., 1961), and as is made clear by the data presented in Table 2.

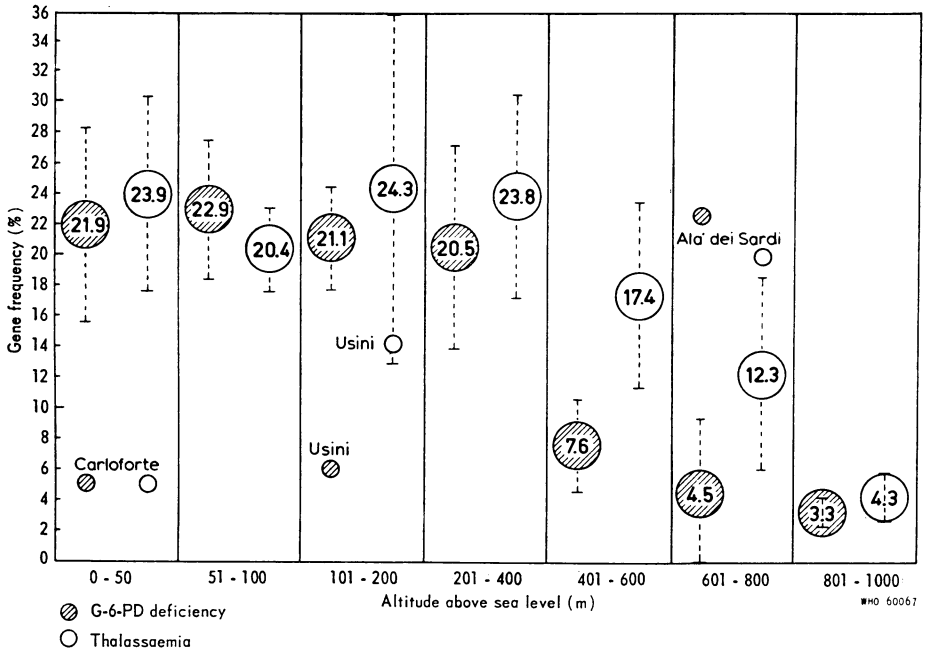
A few villages included in Fig. 2 and 3 (Carloforte, Usini, Lode and Ala dei Sardi) require special mention.

Carloforte is the only village on the beautiful little island of San Pietro, which is adjacent to the very malarial plains of the south-western Sardinian coast. This island was first settled, about AD 1700, by a small group of Genovese fishermen who had come from the island of Tabarca, near the North African coast and, when expelled by the Bey of Tunis, had requested and received the hospitality of the King of Sardinia. This group, now numbering about 7000, kept itself in close isolation from the rest of Sardinia until very recently. It is thus not surprising that, despite the heavy malaria morbidity that was reported in the area until a decade ago, only a few genes for G-6-PD deficiency and thalassaemia can be found among them, and they had clearly been derived from Sardinian ancestors, as could be proved by genealogical studies.

Usini is a small village not far from the north-west coast of Sardinia, where settlements of Genovese and Spanish origin are found. It is in this part of the island that the influence of the Catalan language on the local dialect can be observed most easily. Here, intermixture with the inland population has been more massive and frequent, and the intermediate

FIG. 3

INCIDENCE OF G-6-PD DEFICIENCY AND OF THE THALASSAEMIA TRAIT IN RELATION TO ALTITUDE ABOVE SEA-LEVEL [a]



[a] The figures in each of the large circles are the averages of the gene frequencies found in the villages that fall within the corresponding altitude groupings (0-50 metres, 51-100 metres, etc.). These villages may be found easily in Table 1, in which they are arranged in sequence according to their altitude above sea-level. The smaller circles refer to the three villages of Carloforte, Usini and Ala dei Sardi, which are considered separately, as explained in the text.

TABLE 2. SPLEEN AND PARASITE SURVEYS OF SCHOOLCHILDREN IN 66 SARDINIAN VILLAGES BETWEEN NOVEMBER 1947 AND MARCH 1948, BY VILLAGE ALTITUDE ABOVE SEA-LEVEL [a]

Altitude above sea-level (metres)	Spleen examinations			Parasite examinations					
	Number examined	Number with palpable spleen	Spleen rate (%)	Number examined	Number positive, by *Plasmodium* species			Total number positive	Parasite rate (%)
					vivax	*falciparum*	*malariae*		
0-50	2 858	996	34.8	2 779	59	56	1	116	4.2
51-100	2 015	595	29.5	1 861	39	29	1	69	3.7
101-200	1 282	288	22.5	1 289	58	29	1	88	6.8
201-300	1 215	226	18.6	1 201	26	11	0	37	3.1
301-400	1 094	212	19.4	1 144	12	2	0	14	1.2
401-600	1 540	215	14.0	1 539	12	10	0	22	1.4
601-800	1 481	277	18.7	1 481	18	9	0	27	1.8
801-1000	1 430	106	7.4	1 371]	8	1	0	9	0.7
All altitudes	12 915	2 915	22.6	12 665	232	147	3	382	3.0

[a] Reproduced from Logan, J. A. (1953) *The Sardinian project*, with the permission of the Johns Hopkins Press, Baltimore, Md., USA.

values of gene frequencies today are an obvious consequence. Carcassi (1962) has shown that the same situation prevails for Alghero and its neighbouring villages.

Lode and Ala dei Sardi are two villages that are especially useful to demonstrate the relationship between high gene frequencies and the former prevalence of malaria. Beyond any doubt, these villages originated in the remote past; both of them are noted as sizable settlements in a 16th-century map of Sardinia drawn by Ignazio Donati and now kept in the Vatican Museum. Until a few years ago, their isolation was very strict, since a high mountain, difficult to cross even today, separates them from the coastal villages. Fermi (1938) reported a very high malaria morbidity for both of these villages, despite the relatively high altitude of one of them (Ala dei Sardi); the gene frequencies for both thalassaemia and G-6-PD deficiency are particularly high, unlike those observed in a neighbouring village (Bitti) located on the very summit of the mountain and reported by Fermi (1938) as being relatively free of malaria. All of these villages, together with others in the interior plains, such as Abbasanta and Guspini (see Table 1), again indicate that the suggestion that ethnic heterogeneity is a possible main cause of differences in gene frequencies is certainly not a likely one. Moreover, the blood-group distributions in the interior, coastal and mountain regions are remarkably similar, all showing an unusually high incidence of the M gene and a very low frequency of Rh-negative individuals, which seems to differentiate the Sardinian population from the general European population (Ceppellini, 1955).

INTERACTION BETWEEN G-6-PD DEFICIENCY AND THALASSAEMIA AT THE INDIVIDUAL AND POPULATION LEVELS

As expected, some individuals were found in Sardinia who carry both thalassaemia and G-6-PD deficiency. The association of both of these conditions in the same individual does not appear to involve a more serious red-cell defect, as has been demonstrated directly by chromium-51 studies that have showed that the reduction of red-cell survival time in these individuals is of the same order as that reported for carriers of G-6-PD deficiency alone (Bernini et al., 1964).

Indeed, there are reasons to believe that the association of these two defects in the same person may, rather, produce higher biological fitness in him. For example, we have reported (Siniscalco et al., 1961) that the frequency of severe haemolytic crises from exposure to fava beans (clinical favism) is less among carriers of both the enzyme deficiency and thalassaemia than among carriers of the enzyme deficiency alone. Since G-6-PD activity is always increased in carriers of thalassaemia,[1] it was thought that some kind of compensation for the enzyme deficiency exists in the presence of thalassaemia.

The probability that there is a higher fitness in the carriers of both of these genes is suggested by the finding that the number of such persons in the general population appears to exceed that which would be expected by calculation from the estimated gene frequencies for these traits in each village. While this excess is not significant within any village, it clearly becomes so when the data of the 21 villages studied for that purpose are pooled (Table 3).

OTHER HAEMOGLOBIN VARIANTS AND DIFFERENT FORMS OF THALASSAEMIA IN SARDINIA

To date, there has been no systematic search for haemoglobin variants or for other forms of thalassaemia in Sardinia, although the rarity of haemoglobin variants can be inferred from the scarcity of case reports during the last ten years. On the other hand, the presence of the so-called α-thalassaemia at an appreciable frequency appears probable, from our findings of a certain number of thalassaemic families without elevated haemoglobin A_2 (Carcassi, Ceppellini & Siniscalco, 1957), from the report of Silvestroni & Bianco (1963) of two cases of Bart's haemoglobin among Sardinians living in Rome and from the report of Fiaschi, Campanacci & Naccarato (1964) of a case of thalassaemia-haemoglobin H disease in a haematological patient in the medical clinic of the University of Cagliari.

We have performed an extensive study of nearly 1200 random blood samples collected in villages of the southern plains of Sardinia already known for their high frequency of thalassaemia and G-6-PD deficiency, and not a single instance of variant haemoglobin or of high foetal haemoglobin was found among them. Recently, however, the occur-

[1] Thalassaemic red cells, although smaller than normal ones, have the same enzymatic activity; thus, carriers of the Th (+) gene, who are polycythaemic, have relatively higher G-6-PD activity per blood-volume unit (Piomelli & Siniscalco, 1966, to be published).

TABLE 3

CALCULATION OF THE EXCESS OF DOUBLE CARRIERS IN 21 SARDINIAN VILLAGES

Village No.[b]	Double carriers: Gd(−) & Th(+) [a]				
	Persons tested	Double carriers found	D_i expectation	X_i deviation from expectation	$\frac{X_i}{\sqrt{D_i}}$
2	98	8	7.80	+0.20	+0.072
3	100	11	8.42	+2.58	+0.889
11	100	5	6.76	−1.76	−0.677
14	97	2	2.44	−0.44	−0.281
15	98	3	2.49	+0.51	+0.323
16	100	3	3.68	−0.68	−0.354
24	99	12	7.90	+4.10	+1.459
28	72	4	1.38	+2.62	+2.230
31	96	3	2.63	+0.37	+0.228
33	95	9	9.31	−0.35	−0.100
34	92	3	3.44	−0.44	−0.237
36	98	4	3.13	+0.87	−0.492
37	820	65	55.00	+10.00	+1.350
38	92	2	2.10	−0.10	−0.071
39	99	2	0.24	+1.76	+3.592
40	100	2	1.00	+1.00	+1.000
42	98	2	1.42	+0.58	+0.487
43	100	2	1.54	+0.46	+0.371
44	100	1	1.38	−0.38	−0.323
45	193	3	0.96	+2.0	+2.082
47	80	3	3.24	−0.24	−0.133
					$\sum\left(\frac{X_i}{D_i}\right) = \frac{+14.575}{-2.176}$ / $+12.399$

$$\sum\left(\frac{X_i}{D_i}\right) \Big/ \sqrt{N} \; ; \; \frac{+12.399}{\sqrt{21}} = \frac{12.399}{4.58} = 2.70 \; ; \; P<0.01$$

[a] Gd(−) = glucose-6-phosphate dehydrogenase deficiency; TH(+) = thalassaemia.

[b] See Table 1 for numbering of villages.

rence has been reported of a fast-moving haemoglobin variant that appears to be due to a mutation on the α-haemoglobin chain, probably similar to one that has been described for haemoglobin "Mexico" (Baglioni & Sulis, personal communication). Moreover, an investigation of the distribution of haemoglobin A_2 levels among a random sample of apparent carriers of thalassaemia in an area where this condition is present in about 30% of the population revealed the occurrence of a normal level of haemoglobin A_2 in about 4% of the cases. Further examination of these individuals and of their families led us to the conclusion that they had to be considered instances of α-thalassaemia, although the presence of minor quantities of haemoglobin H could be established in only two of these individuals. If this conclusion is correct, it follows that the incidence of α-thalassaemia in the given area is of the order of 1% ($0.30 \times 0.04 = 0.012$); thus only a minor classification error is involved when popula-

tion-screening for β-thalassaemia is performed by red-cell fragility and blood-film studies alone (Table 4).

The absence of the sickling trait in Sardinia is particularly noteworthy in view of its appreciable frequency in the neighbouring Mediterranean areas (Greece, North Africa, the Middle East, Sicily and, in general, southern Italy) and of its well-established adaptive value in a malarial environment (Allison, 1965).

It has been reported, however, that the incidence of the haemoglobin S gene tends to be correlated inversely with that of β-thalassaemia in those populations in which both of these genes occur with appreciable frequency (Barnicot et al., 1963). This phenomenon has been interpreted as a consequence of the frequently poor adaptability of the genotype that combines haemoglobin S and thalassaemia, thus leading to the elimination of the gene that happened to be the rarest when selective mechanisms of the present type became operative.

Consequently, it may be postulated that, when thalassaemia and G-6-PD deficiency were introduced into Sardinia, the haemoglobin S gene may still have been uncommon in the Mediterranean basin and that it was therefore entirely eliminated from the Sardinian gene pool in the long run, while the other two genes successfully established themselves among the populations of the malarial plains. An alternative explanation is the assumption that both G-6-PD deficiency and thalassaemia are much older mutations than the haemoglobin S gene, and that the enzyme deficiency appeared in the Mediterranean basin when Sardinia had already split off from the mainland and isolated its population.

At any rate, the absence of haemoglobin S from Sardinia is a good piece of evidence for the hypo-

TABLE 4

DISTRIBUTION OF HAEMATOLOGICAL PARAMETERS IN A RANDOM GROUP OF 235 ADULT SARDINIAN MALES (157 NORMALS, 78 WITH THALASSAEMIA) [a]

		Total No.	Means	Standard deviation	Standard errors	Discrimination threshold	Classification error
Haemoglobin	N	157	13.33	1.513	0.121	12.43	27 %
	T	78	11.56	1.460	0.165		
Red cell number	N	157	4.543	0.545	0.043	4.77	34 %
	T	78	5.019	0.601	0.068		
Hematocrit	N	156	42.56	3.571	0.286	40.96	32 %
	T	78	39.17	3.972	0.450		
Mean corpuscular volume	N	156	93.33	9.828	0.787	85.19	21 %
	T	78	77.05	9.823	1.112		
Mean corpuscular haemoglobin	N	157	29.53	3.854	0.308	25.70	16 %
	T	78	22.87	2.854	0.323		
Mean corpuscular haemoglobin concentration	N	156	31.46	2.859	0.229	30.37	35 %
	T	78	29.36	2.629	0.298		
Red cell fragility [b]	N	127	0.413	0.024	0.002	0.378	7 %
	T	61	0.348	0.021	0.003		
Haemoglobin A_2	N	157	2.329	0.392	0.031	3.37	0.7 %
	T	78	5.186	0.680	0.077		

[a] Key: N = normal individuals; T = persons with thalassaemia (i.e., parents of persons with Cooley's disease). Data from Carcassi, Ceppellini & Siniscalco (1957).

[b] Red cell fragility is expressed here as the concentration of NaCl required to produce 50 % haemolysis.

thesis that the Sardinians, unlike their neighbours, must have been genetically isolated for a very long time after the original arrival among them of the genetic raw material upon which natural selection must have been acting for at least 2000 years.

An inverse correlation similar to that described for thalassaemia and sickling appears to exist between the genes for haemoglobins S and C in populations in which both of these variants are present (Allison, 1965). On the other hand, in Greece, a positive correlation has been reported between G-6-PD deficiency and β-thalassaemia (Allison et al., 1963) and, in Greece as well as in Africa, between G-6-PD deficiency and the sickling trait (Motulsky, 1960; Allison, 1965).

These interactions between different genes at a population level are an obvious illustration of the important role that natural selection must evidently play in the maintenance of the genic load of human populations.

POPULATION DYNAMICS OF G-6-PD DEFICIENCY AND THALASSAEMIA

In an attempt to provide a unitary explanation of the world distribution of genes known to involve genetic adaptability in a malarial environment, Zaino (1964) considered the possibility that they may have begun to have an adaptive value more than 50 000 years ago, when Europe was still bridged by land to Africa and malaria was probably already a strong factor in natural selection. When the glaciers receded, the Mediterranean basin was flooded, and its inhabitants either remained isolated along the newly formed seacoasts and islands or migrated towards the Middle and Far East into India, China, and, eventually, the Americas. In their new ecological niches, these genes were exposed to different selective pressures and, under the combined influences of selective pressures such as consanguinity, mutation, interaction between genes, migration, drift, and differential survival, independently reached the diverse equilibrium frequencies that we observe today.

There is naturally no hope of collecting evidence for or against these fascinating arguments. However, it seems to us to be irrelevant to establish whether a set of genes in a given area arose by mutation or migration, as long as it is made clear that their maintenance over an appreciably long time would not be possible without the intervention of selective mechanisms such as those proposed in the classic works of Fisher, Haldane and Wright to explain the occurrence of stable genetic polymorphism.

Livingstone (1964) has recently discussed, in detail, the population dynamics of the genes for thalassaemia, haemoglobin variants and G-6-PD deficiency and has presented the general equations for calculating the time required to attain genetic equilibrium in different selective models. In order to do so, he had to hazard some estimates of the fitness of the various genotypes for each of the genetic systems under consideration. These estimates were probably not far from reality for the haemoglobin S and β-thalassaemia genes, but we cannot agree with those proposed for the G-6-PD deficiency gene, which Livingstone supposed to have a positive adaptive value during malarial periods only in the female heterozygotes. This is, indeed, the obvious conclusion if the gene frequencies observed today are considered " equilibrium frequencies ", but there are no strong reasons to assume that this is necessarily the case. Unlike the situation with the thalassaemia gene, whose lethality in homozygous conditions acts as a strong buffering factor to maintain the system under stable equilibrium at maximum levels of gene frequencies between 0.10 and 0.15, the G-6-PD deficiency gene, which certainly does not produce serious handicaps in its carriers, might well have been in a condition of transient equilibrium in malarial areas. Indeed, we feel that, at least in Sardinia, a situation similar to that presented in Table 5 is closer to reality.

In other words, we assume that, when malaria was killing about one-half of the Sardinian population before reproductive age, the fittest genotypes may well have been the male hemizygote and the female homozygote, despite the slight risk of disease caused by the enzyme deficiency itself, since it is

TABLE 5

THEORETICAL FITNESS VALUES OF THE DIFFERENT GENOTYPES FOR THE THREE SELECTIVE MODELS HYPOTHESIZED IN FIG. 4

	Males		Females		
	G-6-PD deficient	Normals	G-6-PD deficient	G-6-PD inter-mediate	Normals
Model I	1.04	0.96	1.04	1.00	0.96
Model II	1.01	0.95	1.01	1.00	0.95
Model III	0.98	0.94	0.98	1.00	0.94

reasonable to suppose that it might, for instance, have been the susceptibility of the enzyme-deficient red cells of these abnormal individuals to haemolysis that made the growth and multiplication of malarial parasites in their blood more difficult than in the blood of normal individuals. In such a situation, the greater the enzyme deficiency, the greater the protection against malaria.

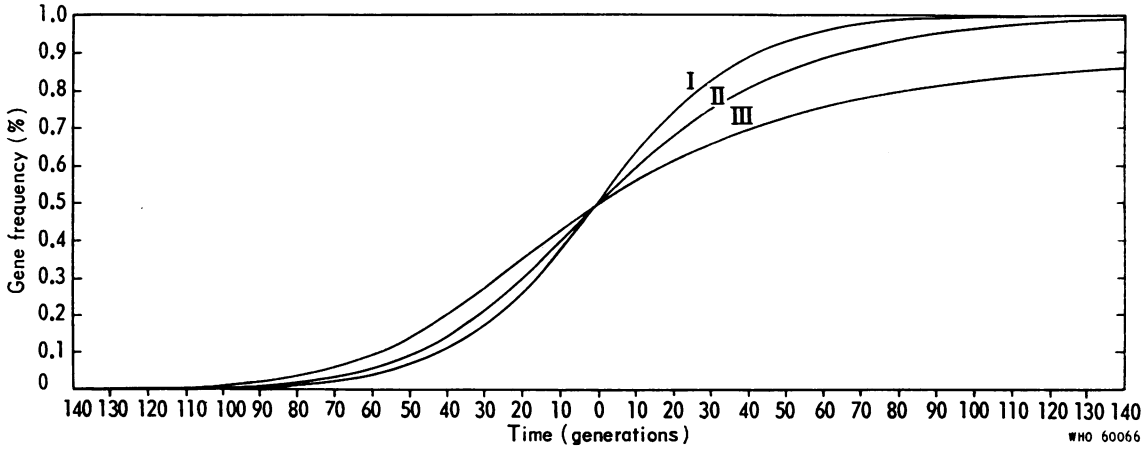
Under such a selective model, the ultimate fate of the enzyme-deficiency gene would, in the long run, have been its complete fixation, had the ecological factor that was responsible for the differential fitness not been removed (Fig. 4). This assumption does not appear to be unduly absurd when one considers the extremely high frequency of this gene in some other populations such as the Iraqi Jews, among whom frequencies of this enzyme deficiency in males range from 25% to 52%, and among the Kurdish Jews, among whom frequencies as high as 70% have been recorded (Szeinberg, 1963).

Nevertheless, it must be borne in mind that such attempts to express, in quantitative terms, the effects of natural selection in human populations are usually over-simplifications, since they can be done only by considering each genetic system separately. There are obvious difficulties in treating, in mathematical terms, more complicated models that would take into account the interactions between various gene systems. This is particularly true when different genes have been selected for by a common ecological factor. In the present instance, it is obviously possible that the greater fitness of the combined genotype Gd(−)Th(+), to which reference was made above, and the total absence of the haemoglobin S gene from Sardinia undoubtedly must have had their weight in influencing the gene-frequency distributions shown in Fig. 2 and 3 and in Table 1.

Moreover, although there are adequate grounds for believing that malaria was the principal ecological factor responsible for the selection of these genes, the possible existence of other genetic and environmental selection factors should not be disregarded. Consanguinity, for example, must have been an important counteracting selective agent for the accumulation of thalassaemia genes in Sardinian isolates, since the increased homozygosity that follows close inbreeding would undoubtedly help in the elimination of the lethal genes. The effects of consanguinity on G-6-PD deficiency must, instead, have been quite the opposite in the presence of malaria, if the fitness estimates shown in Table 5 are correct. To avoid the disturbing effects of consanguinity, we deliberately avoided the inclusion of villages that were of very different sizes and therefore likely to involve significant differences in inbreeding coefficients that were otherwise known to be quite constant in all Sardinian villages of ancient formation and long-standing genetic isolation (A. Moroni, personal communication).

FIG. 4

THEORETICAL CURVES DESCRIBING THE POPULATION DYNAMICS OF THE G-6-PD DEFICIENCY GENE IN SARDINIA UNDER THREE SELECTIVE MODELS [a]



[a] The Roman numerals refer to the three selective models presented in Table 5. The curves were calculated by W. S. Volkers, Department of Human Genetics, University of Leiden, according to the equations proposed by Livingstone (1964).

POSSIBLE GENETIC HETEROGENEITY BETWEEN G-6-PD DEFICIENCY AND THALASSAEMIA PERTINENT TO DIFFERENT AREAS

Unlike the case with haemoglobin variants, for which it can always be established with certainty whether or not one is dealing with a specific mutation (that is, with a specific amino-acid substitution in the haemoglobin molecule), one can never be sure of the genetic homogeneity of the several forms of thalassaemia and G-6-PD deficiency that are reported from different parts of the world. The numerous and not always concordant attempts to classify different forms of thalassaemia (Fessas, 1965) and G-6-PD deficiency (Motulsky, 1965) demonstrate the limitations of our knowledge in this matter, evidently because of the very fact that the diagnoses of thalassaemia and G-6-PD deficiency are made at a level far removed from the primary product of the genes.

For the foregoing reasons, it appeared unwise, at present, to attempt an over-all evaluation of all of the population data so far published. It might well be, for example, that the β-thalassaemia observed in Sardinia is quite different from those reported in Greece or in the Far East, or even from that observed in the Ferrara district. The same is true for G-6-PD deficiency even within the so-called subclasses of the G-6-PD deficiency of the Caucasian type, which involves a complete red-cell enzyme deficiency and no obvious electrophoretic differences, and the G-6-PD deficiency of the African type, which involves a partial red-cell enzyme deficiency associated with not yet well understood changes in the electrophoretic enzyme pattern.

In view of the contrasting conclusions that have been drawn from linkage studies performed in Sardinia and elsewhere to establish the linear sequence on the X chromosome of the G-6-PD gene and other X-borne loci (Siniscalco, 1964; Siniscalco, Filippi & Latte, 1964), we have stressed the possibility that the G-6-PD deficiency may be genetically heterogeneous in different populations. We have also reported some population data that suggest a non-random distribution of colour blindness of the deutan type among enzyme-deficient and normal individuals in Sardinia, suggesting that the neutral or slightly detrimental gene that is responsible for this defect in colour vision has also enjoyed protection in a malarial environment because of its very close linkage with the highly adaptive gene for G-6-PD deficiency (Siniscalco, 1963). Similar observations on the non-random distribution of colour blindness and G-6-PD deficiency have been reported in Israel (Adam, 1963).

If these findings are confirmed, it could mean that, in fact, it is not the G-6-PD deficiency gene alone that confers a higher fitness upon its carriers, but a "successful gene complex" (such as, for example, the G-6-PD deficiency mutant plus a sex-linked modifier that is capable of reducing the red-cell enzyme deficiency to a minimum), which natural selection would have maintained because of its strong adaptive value in a malarial environment. If this were the case, the occurrence in some populations of a genetic mechanism, such as chromosomal inversion, that is capable of reducing crossing-over, could be suspected. This would help to explain how the selective advantage of the G-6-PD deficiency gene may differ from one population to another, if not from one family to another.

CONCLUSIONS

The data that have been reported to date offer striking examples of the possibilities and limitations of studies on human population genetics. On the one hand, we find the availability of true natural populations living in all sorts of ecological situations and therefore exposed to all degrees of natural selection, a detailed knowledge of historical and prehistorical migrations, mating systems, inbreeding, public health and vital statistics, and the facilities for obtaining direct estimates of different genetic fitnesses and exhaustive information concerning their interaction at the individual and population levels. On the other hand, we must consider the difficulty in establishing the relative importance of the numerous identifiable factors of evolution, the time and expense required to collect an amount of data sufficient to justify significant conclusions, the technical difficulties in performing critical studies in areas in which certain forces of natural selection, such as malaria, are still active, and the rapidly increasing impact of modern civilization on human ecology and therefore upon the relative fitness of a given genotype within the time interval of even a single human generation.

However, from all of the foregoing, a conclusion clearly emerges; namely, that it is necessary to learn a great deal more about the genetic structure of our own species and on the biological fitnesses of the principal genotype combinations in the existing "natural environments" as well as in the new environments that our civilization is likely to produce.

ACKNOWLEDGEMENTS

The authors wish to express their gratitude to Professor G. Montalenti, whose encouragement, help and advice were the main factors that made possible the fulfilment of the present research project.

The collection of the blood samples would not have been possible without the generous and unlimited help of the health and administrative authorities of the districts of Cagliari and Nuoro and of the medical officers and school headmasters of the numerous Sardinian villages visited from time to time.

Our gratitude goes also to the members of all of the families investigated, to the Directors of the Ophthalmological Clinic, the Paediatric Clinic, the Institute of Medical Pathology and the Institute of Genetics of the University of Cagliari for their hospitality to the members of the team during the collection of data.

Special thanks go to Mr R. Palmarino, whose thorough and excellent technical assistance was of substantial significance throughout the entire investigation.

The investigations reported in the present paper were supported by grants allocated to the Department of Genetics, University of Rome, by the Rockefeller Foundation (RF 62009, 1962-64), the National Science Foundation (G 19710, 1961-63) and the Italian National Research Council (Commissione di Genetica, 1960-64).

RÉSUMÉ

Les conclusions de nombreux travaux ont appuyé l'hypothèse que le paludisme a représenté un facteur écologique commun responsable de la sélection de trois groupes d'anomalies héréditaires des hématies: hémoglobines atypiques, thalassémie et carence en glucose-6-phosphate déshydrogénase (G-6-PD).

Les auteurs ont étudié la distribution génétique de la thalassémie et de la carence en G-6-PD chez les populations de certains villages sardes dont beaucoup sont très anciens et sont restés isolés pendant de très longues périodes. Ces collectivités offrent les trois conditions permettant une étude des relations possibles entre le paludisme, la thalassémie et la carence en G-6-PD: degré suffisant d'homogénéité ethnique; existence de statistiques démographiques et sanitaires dignes de foi; présence de groupes suffisamment importants exempts de paludisme, dont les caractéristiques ethniques et l'isolement génétique sont semblables à ceux des populations impaludées. Ces situations respectives sont restées inchangées jusqu'à l'achèvement de l'éradication du paludisme en Sardaigne.

Plus de 6000 observations faites dans 52 villages ont mis en évidence une corrélation positive entre l'incidence de la thalassémie et celle de la carence en G-6-PD et la morbidité par le paludisme observée dans le passé. Au cours de l'examen de près de 1200 échantillons de sang, d'autres types d'hémoglobinose (dont l'anémie à hématies falciformes) n'ont pu être décelés.

Les auteurs passent en revue diverses hypothèses concernant le rôle de certaines combinaisons de gènes dans l'adaptation de l'individu à un milieu d'endémie paludéenne. Les porteurs du gène responsable de la déficience enzymatique auraient joui d'une immunité relative vis-à-vis du paludisme; cette action protectrice de la carence en G-6-PD pourrait avoir été influencée diversement par d'autres gènes localisés aux hétérochromosomes.

Les auteurs concluent qu'il est nécessaire de mieux connaître la structure génétique de l'espèce humaine et d'étudier l'aptitude biologique des porteurs des principales combinaisons génotypiques à vivre dans les milieux naturels actuels et futurs.

REFERENCES

Adam, A. (1963) In: Goldschmidt, E., ed., *The genetics of migrant and isolate populations*, Baltimore, Williams & Wilkins, p. 111

Adinolfi, M., Bernini, L., Carcassi, U., Latte, B., Motulsky, A. G. & Siniscalco, M. (1960) *Atti Accad. naz. Lincei R. C.*, **28**, Cl. Sc., 716

Allison, A. C. (1954) *Brit. med. J.*, **1**, 290

Allison, A. C. (1965) *Population genetics of abnormal haemoglobins and glucose-6-phosphate dehydrogenase deficiency*. In: Jonxis, J. H. P., ed., *Abnormal haemoglobins in Africa*, Oxford, Blackwell, p. 365

Allison, A. C. & Clyde, D. F. (1961) *Brit. med. J.*, **1**, 1346

Allison, A., Askonas, B. A., Barnicot, N. A., Blumberg, B. S. & Krimbas, C. (1963) *Ann. hum. Genet.*, **26**, 237

Baglioni, C. (1962) *Correlation between genetics and chemistry of human haemoglobins*. In: Taylor, J. H., ed., *Molecular genetics*, Academic Press, New York. p. 405

Barnicot, N. A., Allison, A. C., Blumberg, B. S., Deliyannis, G., Krimbas, C. & Ballas, A. (1963) *Ann. hum. Genet.*, **26**, 229

Bernini, L., Latte, B., Siniscalco, M., Piomelli, S., Spada, U., Adinolfi, M. & Mollison, P. L. (1964) *Brit. J. Haemat.*, **171**, 10

Brumpt, L. C. & Brumpt, V. (1958) *Bull. Soc. Path. exot.*, **51**, 217

Carcassi, U. (1962) *Giornate di studio sulla microcitemia*, Rome, Istituto Italiano di Medicina Sociale, pp. 78, 182

Carcassi, U., Ceppellini, R. & Pitzus, F. (1957) *Boll. Ist. sieroter. milan.*, **36**, 206

Carcassi, U., Ceppellini, R. & Siniscalco, M. (1957) *Haematologica*, **42**, 1635

Ceppellini, R. (1955) *Cold Spr. Harb. Symp. quant. Biol.*, **20**, 252

Edington, G. M. & Laing, W. N. (1957) *Brit. med. J.*, **2**, 143

Edington, G. M. & Watson-Williams, E. J. (1965) *Sickling, haemoglobin C, glucose-6-phosphate dehydrogenase deficiency and malaria in western Nigeria.* In: Jonxis, J. H. P., ed., *Abnormal haemoglobins in Africa*, Oxford, Blackwell, p. 393

Fermi, E. (1938) *La malaria in Sardegna*, Sassari, Stamperia Libreria italiana e straniera, vol. 2

Fessas, P. (1965) *Forms of thalassaemia.* In: Jonxis, J. H. P., ed., *Abnormal haemoglobins in Africa*, Oxford, Blackwell, p. 71

Fiaschi, E., Campanacci, L. & Naccarato, R. (1964) *Minerva med.*, **55**, 1717

Haldane, J. B. S. (1949) *Ricerca scient.*, suppl. 2, p. 12

Harris, R. & Gilles, H. M. (1961) *Ann. hum. Genet.*, **25**, 199

Ingram, V. I (1963) *The haemoglobins in genetics and evolution*, New York, Columbia University Press

Itano, H. A. (1965) *The synthesis and structure of normal and abnormal hemoglobins.* In: Jonxis, J. H. P., ed., *Abnormal haemoglobins in Africa*, Oxford, Blackwell, p. 3

Kruatzachue, M., Charoeularp, P., Choughsuphajaisiddhim, T. & Harinasuta, C. (1962) *Lancet*, **2**, 1183

Kruatzachue, M., Na-Nakoru, S., Charoeularp, P. & Suwomakul, L. (1961) *Ann. trop. Med. Parasit.*, **5**, 468

Larizza, P., Ventura, S., Mattioli, G., Sulis, E. & Aresu, G. (1958) *Haematologica*, **43**, 517

Livingstone, F. B. (1964) *Am. J. hum. Genet.*, **16**, 435

Logan, J. A. (1953) *The Sardinian project*, Baltimore, Johns Hopkins Press

Motulsky, A. G. (1960) *Hum. Biol.*, **32**, 28

Motulsky, A. G. (1965) *Theoretical and clinical problems of glucose-6-phosphate dehydrogenase deficiency: its occurrence in Africans and its combination with haemoglobinopathy.* In: Jonxis, J. H. P., ed., *Abnormal haemoglobins in Africa*, Oxford, Blackwell, p. 143

Pérez, N., Clark-Turri, L., Rabjille, E. & Niemeyer, H. (1964) *J. biol. Chem.*, **239**, 2420

Rucknagel, D. L. (1964) *Ann. N.Y. Acad. Sci.*, **119**, 436

Rucknagel, D. L. & Neel, J. V. (1961) *The hemoglobinopathies.* In: Steinberg, A. G., ed., *Progress in medical genetics*, New York, Grune & Stratton, p. 158

Silvestroni, E. & Bianco, I. (1963) *Le emoglobine umane*, In: Gedda, L., ed. *De genetica medica*, Rome, Istituto Mendel, vol. 4, pp. 197-336

Siniscalco, M. (1963) *Linkage data for G6PD deficiency in Sardinian villages.* In: Goldschmidt, E., ed., *The genetics of migrant and isolate populations*, Baltimore, Williams & Wilkins, p. 106

Siniscalco, M. (1964) In: Geerts, S., ed., *Genetics today. Proceedings of the XIth International Congress of Genetics, The Hague, 1963*, London, Pergamon Press, vol. 2, p. 851

Siniscalco, M., Bernini, L., Latte, B. & Motulsky, A. G. (1961) *Nature (Lond.)*, **190**, 1179

Siniscalco, M., Filippi, G. & Latte, B. (1964) *Nature (Lond.)*, **204**, 1062

Szeinberg, A. (1963) *G6PD deficiency among Jews—genetic and anthropological considerations.* In: Goldschmidt, E., ed., *The genetics of migrant and isolate populations*, Baltimore, Williams & Wilkins, p. 69

Thompson, G. R. (1962) *Brit. med. J.*, **1**, 682

Zaino, E. C. (1964) *Ann. N.Y. Acad. Sci.*, **119**, 404