PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Genetic Factors as Determinants of Infectious Disease Transmission in Human Communities [and Discussion]

D. J. Weatherall, J. I. Bell, J. B. Clegg, J. Flint, D. R. Higgs, A. V. S. Hill, G. Pasvol, S. L. Thein and C. C. Blackwell

Phil. Trans. R. Soc. Lond. B 1988 **321**, 327-348
doi: 10.1098/rstb.1988.0095

Phil. Trans. R. Soc. Lond. B **321**, 327–348 (1988)
Printed in Great Britain

Genetic factors as determinants of infectious disease transmission in human communities

By D. J. Weatherall, F.R.S., J. I. Bell, J. B. Clegg, J. Flint, D. R. Higgs, A. V. S. Hill, G. Pasvol and S. L. Thein

MRC Molecular Haematology Unit, Nuffield Department of Clinical Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, U.K.

Genetic factors may play an important role in individual susceptibility to infection. Hitherto this problem has been investigated by attempting to relate the distribution of genetic polymorphisms in populations to present or past infection, or by analysing specific infections by classical twin studies or group comparisons. There is reasonable evidence that the common red-cell polymorphisms involving haemoglobin, enzymes or membrane have been maintained by relative resistance to malaria. Blood-group heterogeneity, including secretor status, may reflect varying susceptibility to bacterial, virus and yeast infection. There is increasing evidence that the HLA-DR system may be involved in modifying the clinical course of bacterial, virus and parasitic infection. So far no specific resistance or susceptibility loci similar to those found in murine models have been found in man. DNA analysis, particularly involving restriction fragment length polymorphism associations with candidate genes, offers a valuable new approach to this problem.

Introduction

It has long been suspected that genetic factors play an important role in determining susceptibility to infection. While this subject is of considerable biological interest it may also have important practical implications. Following the advent of recombinant DNA and the new techniques of protein engineering, many candidate vaccines are becoming available that will need to be evaluated in defined populations. As it is quite possible that at least some of them, although useful, will not afford complete protection, any trials designed to test their efficacy will have to take into account the possibility that, because of their genetic makeup, some individuals will be more or less susceptible to a particular organism than the rest of the population. Furthermore, it is possible that there may be inherited variability in response to vaccines.

The mechanisms for protection against infection are extremely complex, ranging from the natural dermal and mucosal barriers, through the activity of the humoral immune system, to extremely sophisticated mechanisms for cellular killing mediated through complement, polymorphonuclear leucocytes and macrophages, and lymphocytes. Clearly, many genes must be involved in this diverse protective armamentarium, the variation of which might modify susceptibility to infection. But with the exception of a few rare mutations that cause gross defects in resistance to infection very little is known about the genetic variability of these systems that might play a major role in modifying the pattern of infectious disease in large populations.

Much of our extremely limited understanding about genetic factors in susceptibility to infection has come from the analysis of the relative distributions of genetic polymorphisms in

relation to present or previous infection in small groups or large population samples. For example, soon after it was found that there is a wide diversity in the distribution of blood groups between different races it was suggested that this might reflect variability in response to infectious disease in individuals of different genetic makeup. However, it turned out to be extremely difficult to test this hypothesis and, as we shall see later, with a few exceptions there is very little hard evidence that relates a particular blood group to susceptibility or resistance to infection. However, the brilliant insight of J. B. S. Haldane, who in 1949 suggested that heterozygote advantage to malaria might be the mechanism for the very high gene frequency of thalassaemia in the tropical world, opened up a field of investigation that has, to date, provided the best model for studying the population genetics of resistance to infection. Although there has been more recent work on other systems, particularly the human lymphocyte antigens (HLA-DR) genes, the genetic disorders of the red cell remain the only polymorphisms for which we are even close to understanding the complexities of host–parasite interaction.

In this paper we shall review current knowledge about the association of genetic polymorphism and infection and, because they may provide us with some clues for future directions for work in this field, summarize briefly the single-gene disorders that are known to cause increased susceptibility to infection.

Genetic polymorphisms and infection

Differences between ethnic groups, or between subsets of individuals within particular races, can be described in terms of measurements and comparisons of gene frequencies for polymorphic discontinuous traits. Such differences can in theory result from mutation, hybridization, genetic drift and selection. Thus in any study designed to test the hypothesis that a particular distribution of polymorphisms is due to selection by resistance or increased susceptibility to an infective agent, it is essential to consider all these factors. In practice it is particularly important to distinguish between adaptive and non-adaptive divergence, that is between selection and drift. This means that the two populations being compared must be as alike as possible in every respect except in their exposure to the particular infective organism under examination. As we shall see, many studies in this field are bedevilled with this particular difficulty.

In attempting to detect selection by infectious disease as the basis for different gene frequencies a number of factors must be considered. Assuming random mating within a population, the frequencies of genotypes are easily explained as the product of the frequencies of the genes corresponding to the genotypes. This is the basis of the Hardy–Weinberg distribution; a homozygote for an allele will have a frequency equal to that of the square of the frequency of that allele, and heterozygote frequency will be twice the product of the two corresponding allele frequencies. Deviations from this pattern, if they are not due to non-random mating or population heterogeneity, may indicate the existence of selective factors. For example, in some of the red-cell polymorphisms that we shall consider it is possible to compute fitness directly from comparisons of observed genotype frequencies with their expected values; these direct estimates of fitness lead to expected gene frequencies that are very close to those observed.

Some of the polymorphic systems that we shall consider are the products of families of linked

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

genes. In the absence of selection, and given sufficient time for equilibrium to have been reached, no associations between alleles due to linkage disequilibrium are to be expected. If they occur it suggests that selection may be favouring a particular linkage haplotype. Again, however, linkage disequilibrium is only evidence of a selective process if it can be established with certainty that there has been no recent hybridization, that is population admixture.

Table 1 summarizes some of the polymorphic systems that have been examined for a possible relationship to susceptibility to infectious disease. The polymorphisms involving the red cell are by far the best studied. They include haemoglobin, enzyme systems and membrane structure. The genetic haemoglobin variants that achieve polymorphic frequencies include haemoglobins with altered structure, notably haemoglobins S, C and E, and the thalassaemias, conditions that result from a reduced rate of production of either the α or β globin chains of adult haemoglobin. The only red-cell enzyme defect that reaches polymorphic frequency is glucose-6-phosphate dehydrogenase (G6PD) deficiency. This enzyme plays a key role in the hexose monophosphate shunt pathway of red cells that is essential for their protection against oxidant damage. Finally, a red-cell membrane defect of unknown cause, hereditary ovalocytosis, reaches polymorphic frequencies in a few localized populations, notably those of Papua New Guinea. As we shall see later there is reasonably good evidence that all these polymorphisms have been maintained because of the resistance that they afford against *Plasmodium falciparum* malaria.

TABLE 1. SOME POLYMORPHIC SYSTEMS ASSOCIATED WITH INFECTIOUS DISEASE

human
red cell
haemoglobin
enzymes
membrane proteins
blood groups
HLA-DR
class I gene products
class II gene products
immunoglobulin
Gm
Km
complement
mouse
MHC
Lsh
Ity
Mcp
others

The HLA complex is a multi-gene family on the short arm of chromosome 6 that codes for molecules that are critical for self–nonself discrimination. The gene products can be divided into two main types. First, there are the class I molecules that are determined by the HLA-A, B and C genes. They consist of a heavy chain that is non-covalently associated with a 12000 kDa chain, β_2 microglobulin. The HLA-A, B and C genes code for class I heavy chains; β_2 microglobulin is encoded separately by a locus on chromosome 15. The class I molecules are involved in target cell recognition by a subset of T cells with suppressor or cytotoxic activity. They are particularly important in viral antigen presentation and in the subsequent

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

recognition of infected cells by cytotoxic T cells. Second, there is the DR region that encodes for class II gene products that appear on cell surfaces as heterodimers. The structural genes for both chains of these molecules, α and β, are all located within the HLA-DR region. Three distinct families of class II molecules have been characterized: DR, DQ and DP. They are also essential for antigen recognition by T cells, in this case of the helper variety.

The genes of the HLA complex are highly polymorphic, judged both serologically and, more recently, by DNA analysis using restriction fragment length polymorphisms (RFLPS). A considerable amount of work has been carried out in an attempt to relate specific HLA haplotypes with diseases, particularly of the autoimmune variety. As we shall see later, many of the associations with infectious disease are extremely weak and the best evidence for the importance of this gene cluster in modifying response to infection comes from studies of conditions such as leprosy in which it appears that particular polymorphisms are involved with modification of the course of the illness.

Finally, there are a number of other highly polymorphic systems, particularly the blood groups, for which a relationship to infectious disease is suspected but where, with a few notable exceptions, hard experimental evidence is lacking.

Genetic disorders of the red cell and the malaria hypothesis

In 1949 Haldane suggested that heterozygotes for thalassaemia might be protected against malaria and that this might be an important mechanism whereby these conditions have attained their extraordinarily high gene frequencies. In other words the world distribution of these disorders might reflect a state of balanced polymorphism. E. B. Ford has defined polymorphism as 'the occurrence together in the same habitat at the same time of two or more distinct forms of the species in such proportions that the rarest of them cannot be maintained merely by recurrent mutation'. The term 'balanced polymorphism' implies that the gene frequencies for the advantageous heterozygous states for conditions such as thalassaemia will increase until they are balanced by the loss of homozygotes from the population.

Haldane made these predictions before it was known that thalassaemia is heterogeneous or that, as well as sickle-cell anaemia, there are a number of other haemoglobin variants, notably haemoglobins C and E, that reach polymorphic frequencies (table 2). Furthermore, he knew nothing of G6PD deficiency or ovalocytosis. In the sections that follow we shall ask how well the 'malaria hypothesis' has stood the test of time and whether it is justified to extend it to cover these other common red-cell polymorphisms.

The geographical distribution of the red-cell polymorphisms

The red-cell polymorphisms only occur at high frequencies in regions in which malaria has been common in the past or is still endemic (WHO 1982).

TABLE 2. RED-CELL POLYMORPHISMS ASSOCIATED WITH MALARIA RESISTANCE

P. falciparum
haemoglobins S, C and E
α thalassaemia
β thalassaemia
G6PD deficiency
ovalocytosis
P. vivax
Duffy blood group

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The sickle-cell gene is distributed in a broad band across tropical Africa and is found at lower frequencies in some non-African Mediterranean populations. It also occurs throughout the Middle East and central India, but apart from some populations of Indian origin in north Malaya it has not been observed in Southeast Asia. Haemoglobin C is restricted to parts of West Africa whereas haemoglobin E is distributed in a region stretching from the eastern parts of India through Burma to Southeast Asia where it reaches its highest frequency in parts of Thailand, Laos and Cambodia. It occurs sporadically in parts of southern China and it extends south in a line stretching through Thailand down the Malay peninsula into some of the island populations of Indonesia.

Approximate gene frequencies for these structural haemoglobin variants have been catalogued by several workers (Livingstone 1985; Bunn & Forget 1986). In regions of West Africa and parts of eastern Saudi Arabia and central India the gene frequency for haemoglobin S can exceed 0.15. Similar frequencies for haemoglobin C are found in Ghana and Upper Volta. The gene frequency for haemoglobin E in Thailand and Burma is 0.05–0.10, although higher values have been recorded in parts of eastern Thailand near the Vietnamese border in the so-called 'haemoglobin E triangle'.

The world distribution of the thalassaemias, together with their underlying mutations, is summarized in figures 1 and 2. Adequate gene frequency data for β thalassaemia (figure 1) are only available for a limited number of populations (Weatherall & Clegg 1981; Livingstone 1985). The condition is uncommon throughout tropical Africa although high frequencies have been observed in north Africa. It is particularly common throughout the Meditteranean basin, parts of the Middle East, and in India and Burma and it extends thoughout Southeast Asia in a line starting in southern China, stretching through Thailand, Cambodia and Laos and down the Malay peninsula into some of the island populations.

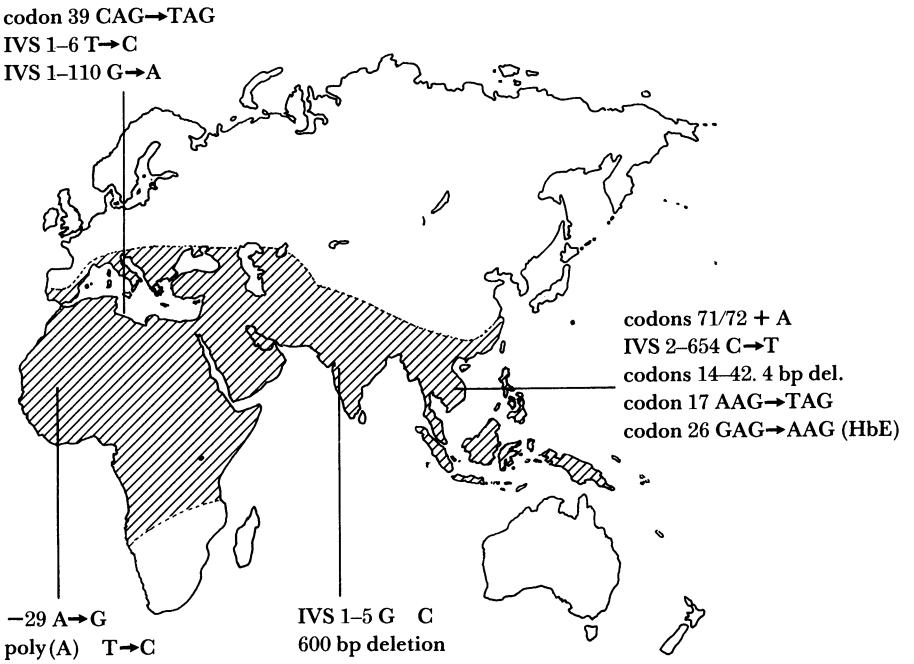

FIGURE 1. The world distribution and main different mutations of the β thalassaemias.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

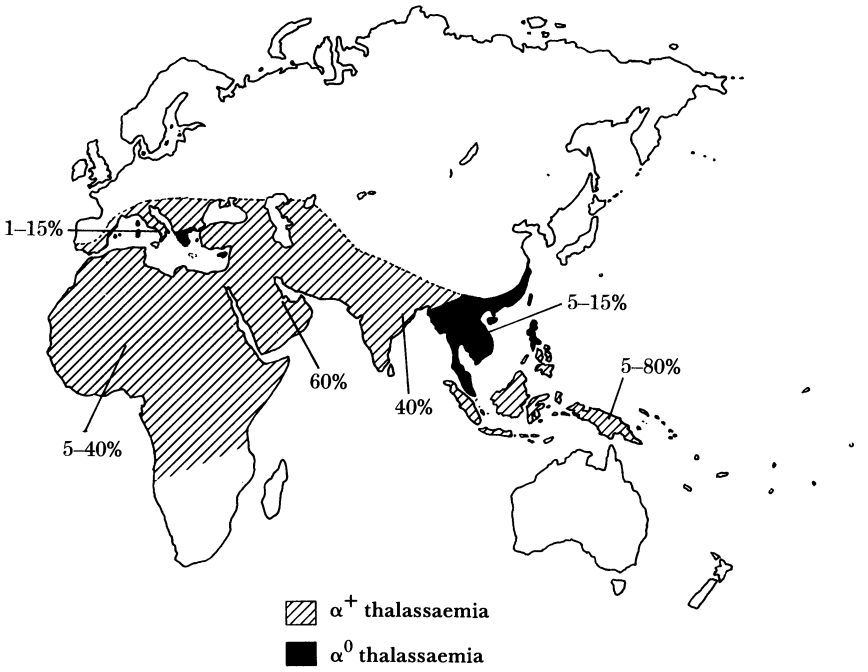

FIGURE 2. The world distribution of the α^0 and α^+ thalassaemias with very approximate gene frequencies in different populations.

The world distribution of the α thalassaemias is summarized in figure 2. The α^+ thalassaemias, that is those due to loss of activity of one of the linked α globin genes on chromosome 16, are extremely common in parts of Africa, the Mediterranean region, the Middle East and throughout Southeast Asia and the Pacific island populations. Indeed, α^+ thalassaemia seems to be approaching fixation in some regions, notably the coastal regions of Papua New Guinea (Oppenheimer *et al.* 1984; Flint *et al.* 1986). On the other hand, the α^0 thalassaemias, which are due to deletions of both α globin genes, are restricted to the Mediterranean island populations and to Southeast Asia where they occur most frequently in China, Thailand, Laos, and the Malay peninsula; α^0 thalassaemia is not found in the island populations of Melanesia (Higgs & Weatherall 1983; Weatherall *et al.* 1988).

By and large, G6PD deficiency follows a similar distribution to the haemoglobin variants. It occurs at a high frequency throughout tropical Africa, the Mediterranean region, the Middle East and Southeast Asia (Livingstone 1985). Ovalocytosis is a common dominant disorder that is found sporadically in every racial group and is undoubtedly heterogeneous at the molecular level. However, there appears to be a high frequency of a mild form in Papua New Guinea (Kidson *et al.* 1981).

Origins and movements of the genes for the red-cell polymorphisms

The remarkably high gene frequencies for the common red-cell polymorphisms have given rise to much speculation about their origin and how they have become distributed among different populations. Until recently it was believed that many of them had arisen as the result of one mutation and that they were then subjected to a combination of selection and drift due to population movements. For example, it was believed that the distribution of the sickle-cell gene in the New World reflects emigration from West Africa during the transportation of

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

slaves. Similarly, it was thought that its high prevalence in Saudi Arabia and India may have resulted from migrations out of East Africa; the transport of slaves from East Africa to the Persian Gulf flourished from 200–1500 A.D. Based on similar reasoning, it was thought that the β thalassaemia gene had moved eastwards into Southeast Asia from an origin in the Mediterranean basin (see Weatherall & Clegg 1981; Bunn & Forget 1986).

The application of recombinant-DNA technology to the study of human globin genes has made it necessary to re-examine many of these ideas. Two discoveries in particular have led to this reappraisal. First, it has been found that all forms of thalassaemia show remarkable molecular heterogeneity and that most populations with a high frequency of the disorder have their own unique mutations (figure 1). Secondly, it has been possible to start to analyse the evolutionary history of the globin-gene mutations by examining the pattern of restriction fragment length polymorphisms (RFLPs) in and around the affected α and β globin genes. An RFLP is a harmless genetic variation in which a single base change in DNA either creates a new site for cleavage by a restriction endonuclease or removes a previously existing one; the result is an inherited difference in the size of the DNA fragments generated by the particular enzyme.

It has been found that the pattern of RFLPs in the human globin-gene clusters is not random but rather that there is a limited number of arrangements, or haplotypes, in different populations (Antonarakis *et al.* 1982; Wainscoat *et al.* 1986). Current evidence suggests that the haemoglobin mutations occurred after these haplotype patterns were laid down in different populations and that the haplotypes that were associated with common globin-gene mutations were then selected along with them.

The first study relating RFLPs to the distribution of haemoglobin variants concerned the sickle-cell mutation and a single polymorphism near the β globin gene that is identified by the restriction enzyme *Hpa*I (Kan & Dozy 1980). Nearly all Caucasians and about 90% of Africans with a normal haemoglobin phenotype have a 7.6 kilobase (kb) or, less commonly, a 7.0 kb *Hpa*I fragment that encompasses the entire β globin gene; only 3% have a 13 kb fragment. In contrast, among American blacks who have haemoglobin S, nearly 70% have the 13 kb fragment. Similarly, blacks with haemoglobin C nearly all have the 13 kb fragment. It seems likely, therefore, that the β^S and β^C mutations both arose on a β gene carrying a 13 kb polymorphism that may have been restricted to a relatively small geographical area corresponding to what is now called Burkina Faso, and Ghana. Presumably because the β^S and β^C genes came under strong selection their frequency increased and the linked 13 kb polymorphism 'hitch-hiked' along with them. On the the other hand the prevalence of the ancestral 13 kb β^A gene remained low because it lacked any selective advantage. The association between the β^S mutation and the 13 kb polymorphism has also been found in north Africans and Sicilians although not in individuals from East Africa, Saudi Arabia or India. These observations suggested that the sickle-cell gene arose independently in West and East Africa (Mears & Lachman 1981).

Recently, extensive RFLP analyses using a battery of single point polymorphisms in the β globin gene cluster have been done in several populations in an attempt to define further the origins of both the sickle-cell and haemoglobin E mutations. Studies in Jamaica and West Africa have shown that the β^S mutation can be found in association with a wide variety of haplotypes (Wainscoat *et al.* 1983; Antonarakis *et al.* 1984). Although these results are compatible with multiple origins for the β^S mutation they must be interpreted with caution

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

because it is possible that a number of gene conversion events could have given rise to at least some of the variability of the β^S–haplotype associations. However, the occurrence of a common haplotype associated with the β^S gene in parts of eastern Saudi Arabia and Orissa in central India, and the finding of completely different haplotypes in some African populations, suggests that the β^S mutation had at least two origins, one in Africa and another in the Middle East or India (Kulozik *et al.* 1986, 1987) (figure 3). Similar arguments are applicable to the β^E mutation, but although it has been found with several different haplotypes it is not absolutely certain whether this reflects multiple origins or the action of a number of gene conversions (Kazazian *et al.* 1984).

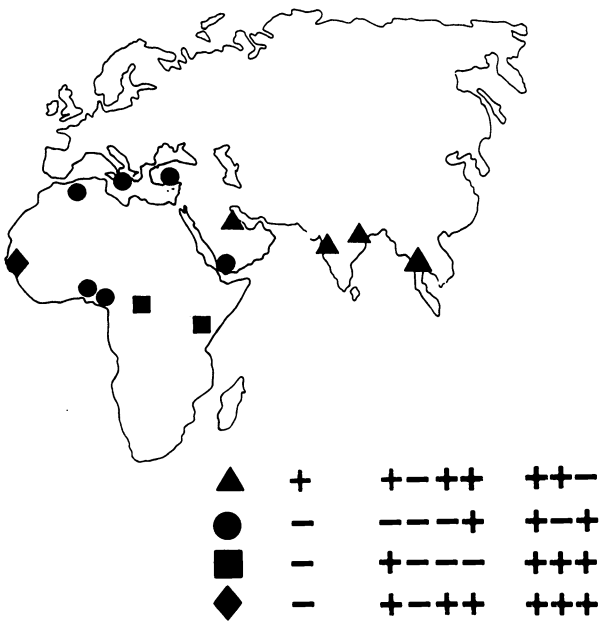

FIGURE 3. The β globin gene restriction fragment length polymorphism haplotypes in different sickling populations. The + and − designations represent the following polymorphic sites in the β globin gene cluster (left to right): *Hin*dII 5′ to the ε gene, *Hin*dIII in the $^G\gamma$, *Hin*dIII in the $^A\gamma$ gene, *Hin*dII in the ψβ gene, *Hin*dIII downstream from the ψβ gene, *Ava*II in the β gene, *Hpa*I downstream from the β gene, and *Bam*H1 downstream from the β gene, respectively. Based on Kulozik *et al.* (1986).

It is now apparent that in each of the high-frequency areas there are a few common β thalassaemia mutations together with a varying number of rare ones (Orkin & Kazazian 1984; Weatherall *et al.* 1988) (figure 1). Furthermore, in each of these regions the mutations are different (Weatherall & Clegg 1981; Bunn & Forget 1986; Weatherall *et al.* 1988). And even if the same mutation does occur in different populations it is nearly always found in association with a different β globin gene RFLP haplotype. These findings suggest that the β thalassaemia mutations have arisen independently in different populations and then achieved their high frequency by selection. Although there may have been some drift and population movement of the β thalassaemia genes between populations there is now no doubt that independent mutation and selection provides the best explanation for their world distribution.

In general, recent studies of the α thalasaemias lead to similar conclusions. The deletions that cause α^0 thalassaemia in the Mediterranean region and Southeast Asia are different (Pressley *et al.* 1980; Higgs & Weatherall 1983). Furthermore, studies of the RFLP haplotypes in the α

globin gene cluster, together with a careful analysis of the crossover events that have produced the α thalassaemias, have shown that the common form of α^+ thalassaemia due to a 3.7 kb deletion has occurred more than once in different populations. Furthermore, there are less common forms of α^+ thalassaemia that are confined to particular populations, the Pacific islanders of Melanesia and Polynesia for example (Higgs *et al.* 1984; Hill *et al.* 1985). Haplotype analyses suggest that the α^+ thalassaemias found in Melanesia arose there *de novo* and reached their high frequencies by selection; the haplotypes of the α^+ thalassaemias in Southeast Asia are completely different (Flint *et al.* 1986). Thus it is likely that the different types of α^+ thalassaemia have arisen independently in many different populations and, like the β thalassaemias, have reached their very high frequencies by selection.

Similar conclusions can be drawn about the distribution and heterogeneity of G6PD deficiency (Luzzatto 1985). There are numerous different G6PD variants, most of which tend to segregate within particular racial groups. Thus it seems likely that they also arose independently and reached their high frequencies by selection.

The world distribution of the common red-cell variants and their association with present or past malaria

The overall world distribution of the common haemoglobin variants and thalassemias tends to mirror that of malaria. By and large all the haemoglobin disorders are distributed among present or past malarious areas, although by no means evenly. Thus whereas α^+ thalassaemia occurs at a high frequency in populations where the sickle-cell, haemoglobin E or β thalassaemia genes are found, it is unusual to encounter the coexistence of more than one β globin gene variant at high frequencies in the same population. The more patchy world distribution of the latter probably reflects the relative fitness of the different heterozygous states and the deleterious effects of the compound heterozygous states for β thalassaemia and the β chain haemoglobin variants. For example, in Africa the sickle-cell gene has been successful at the expense of β thalassaemia. Presumably selection has been stronger for the sickle-cell trait; the compound heterozygous state for β thalassaemia and haemoglobin S, sickle-cell thalassaemia, has a fairly severe clinical phenotype and so β thalassaemia has not reached a high freqency. Presumably the same mechanism has been at work in Southeast Asia where the haemoglobin E gene seems to have prospered over β thalassaemia, although here the situation is not so clear-cut because β thalassaemia is still quite common. Because of the relative lack of interaction between α thalassaemia and the different β chain haemoglobin variants, α thalassaemia has come under strong selection and is common in all these populations. One curious anomaly is the fact that whereas α thalassaemia is widely disseminated throughout Melanesia and Polynesia there are no common β structural variants in this part of the world and β thalassaemia occurs only sporadically in a few island isolates (Hill *et al.* 1988).

Early epidemiological studies in Africa suggested that haemoglobin S heterozygotes are protected against *P. falciparum* malaria (Allison 1954). More recent work in Nigeria has confirmed these studies (Fleming *et al.* 1979). For example, the prevalence of the sickle-cell trait in Nigerian newborns is 24% compared with 29% over the age of five; above this age no differences are found. These observations suggest that the sickle-cell trait confers a relative fitness of about 0.20 compared with normal individuals in the same population. This figure gives a gene frequency of about 0.15, a value very close to that observed in Nigeria. In addition, it can be calculated that it would take 50 generations or about 1000 years to reach the present equilibrium assuming that the homozygous condition is 100% lethal. These studies have been

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

supported by analysis of the frequency of malarial parasites in the blood of sickle-cell heterozygotes of different ages; the relative numbers were reduced in carriers aged 30–60 weeks but not in older children.

Population data of this kind relating β thalassaemia heterozygosity to malaria, although suggestive, have provided less clear-cut information (see Weatherall & Clegg 1981). Studies in Sardinia, which showed that β thalassaemia is less common in the mountainous regions where malarial transmission was relatively low, suggested that β thalassaemia might have reached its high frequency owing to protection against malarial infection (Siniscalco *et al.* 1966). These conclusions were questioned by Brown (1981), who suggested that much of the observed altitude differences in distribution of β thalassaemia might reflect gene flow rather than selection.

Regardless of these criticisms, for many years the Sardinian data remained the only convincing evidence for the protective effect of thalassaemia against malaria, despite many attempts to study this problem (Weatherall & Clegg 1981). However, recent work utilizing malaria endemicity data and globin-gene mapping has shown a very clear altitude-related effect on the frequency of α thalassaemia in Papua New Guinea (Oppenheimer *et al.* 1984; Flint *et al.* 1986). Furthermore, a sharp cline in the frequency of α thalassaemia has been found in a zone stretching south from Papua New Guinea through the island populations of Melanesia to New Caledonia. A similar gradient in the distribution of malaria has been demonstrated from previous spleen-rate data collected in this region over many years. This observation might also have reflected gene drift and founder effects in these island populations. For example, a population with a high frequency of α thalassaemia might have moved through the islands from the north and the gene frequency have been diluted during their migration southwards. This explanation is unlikely, however, because there is a random pattern of the distribution of other DNA polymorphisms throughout these island populations (Flint *et al.* 1986; Hill 1986). Thus the frequency of α^+ thalassaemia, but not that of other DNA polymorphisms, shows a clear altitude- and latitude-dependent correlation with malarial endemicity throughout Melanesia.

Cellular mechanisms for protection against P. falciparum malaria

Despite a considerable amount of work the cellular mechanisms that might mediate resistance of genetic red-cell variants to malarial parasites remain uncertain. The problem has been approached by *in vitro* studies of the invasion of *P. falciparum* in either normal red cells or those that contain haemoglobin variants or that are G6PD deficient.

In the case of the sickle-cell trait, experiments have been designed to study the rates of sickling of parasitized and non-parasitized cells under reduced oxygen tension and the patterns of invasion and growth of parasites in red cells containing haemoglobin S in *in vitro* cultures. Two potential protective mechanisms have been demonstrated. First, it has been found that under low oxygen tension sickle-cell trait cells containing *P. falciparum* sickle more readily than non-parasitized cells (Luzzatto *et al.* 1970). It has been suggested, therefore, that accelerated *in vivo* sickling might provide a mechanism for the rapid clearance of parasitized cells and their subsequent destruction in the reticulo-endothelial system. This would ensure that the parasite would not complete its life cycle. *In vitro* culture studies have also shown that, under ambient oxygen tensions, parasites invade and grow in sickle cells at the same rate as normal cells. However, under reduced oxygen conditions both invasion and growth is inhibited (Friedman

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

1978; Pasvol & Weatherall 1978). It has been suggested that this effect may be mediated at least in part by the relatively low levels of potassium in AS cells that have undergone sickling (Friedman *et al.* 1979).

It is conceivable that both mechanisms are involved; the major reason why the entrapment of sickle cells has been favoured is the observation that individuals who are homozygous for sickle-cell anaemia may undergo severe fatal malarial infection, an observation that is difficult to equate with the lack of survival of parasites in cells that contain large concentrations of haemoglobin S. On the other hand, as even a mild malarial infection might have catastrophic consequences for a child with sickle-cell anaemia, this objection may not be valid.

The results of *in vitro* studies of parasite invasion and growth in heterozygous thalassaemic cells are inconsistent. Some have failed to show any convincing evidence that these cells are more resistant to malarial infection (Pasvol & Wilson 1982) whereas others have demonstrated some degree of inhibition of invasion (Brockelman *et al.* 1987). It has been found that cells containing relatively large amounts of haemoglobin F tend to show a reduced rate of parasite growth (Pasvol *et al.* 1977) and it is known that the rate of decline of haemoglobin F production over the first year of life is retarded in β thalassaemia heterozygotes (Weatherall & Clegg 1981). On the other hand, in most studies the red cells of both α and β thalassaemia heterozygotes are invaded at the same rate as normal cells by parasites in culture and the patterns of growth and development show no difference. *In vitro* experiments have shown that the red-cell membrane in heterozygous thalassaemia is particularly sensitive to damage by oxidation and that infection with the malarial parasite may produce sufficient oxidative stress to alter intracellular metabolism in the way that might lead to the premature death of the parasite (Friedman 1979; Friedman & Träger 1981). However, these experiments were carried out under such extreme conditions of stress that their *in vivo* relevance is unclear.

The proposed mechanisms whereby G6PD deficient cells may be resistant to *P. falciparum* malaria are even more complex (Luzzatto 1985). It has been found that these cells, maintained in standard culture (Roth *et al.* 1983), or in conditions of oxidative stress (Friedman 1979), support growth of *P. falciparum* at a reduced rate as compared with normal cells. However, because this is an X-linked disorder, *in vivo* protection must be a prerogative of G6PD-deficient heterozygous females and not of deficient hemizygous males (Luzzatto 1985). It has been found that G6PD deficient cells become as competent as normal cells in supporting the growth of malarial parasites if the latter have already undergone several cycles of schizogony (Luzzatto *et al.* 1983). This suggests that an adaptive process is taking place in the parasite, by using its own G6PD gene for example, and provides a possible mechanism whereby, *in vivo*, heterozygotes are relatively resistant to *P. falciparum* malaria whereas hemizygotes are not. In a hemizygous male with only G6PD-deficient cells such adaptation will take place rapidly and parasitaemia may reach dangerous levels in the same way as in a normal individual. On the other hand in a heterozygous female, in whom there are both normal and deficient red-cell populations, the parasite would have an equal chance of entering deficient or normal cells, thus decreasing the chance for adaptation. Other mechanisms for resistance of G6PD-deficient cells to malaria have been suggested, including altered thiol status (Miller *et al.* 1984).

The red cells of patients with the form of ovalocytosis found commonly in Papua New Guinea are relatively resistant to invasion by *P. falciparum in vitro* (Kidson *et al.* 1981). This may be related to their relative rigidity due to an abnormality of their cytoskeleton.

In vitro studies of the rates of invasion and growth of *P. falciparum* in red cells containing

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

haemoglobin E have given inconclusive results; this variant produces the phenotype of a mild form of β thalassaemia and hence the same protective mechanisms may be involved in this common condition as in the different forms of thalassaemia (Pasvol & Wilson 1982; Luzzatto 1985). Red cells from haemoglobin C homozygotes, but not heterozygotes, show reduced parasite growth (Pasvol & Wilson 1982). Again the mechanism is not clear.

Could there be a factor that most of the red-cell variants have in common that makes them more resistant to invasion by *P. falciparum*? In this context it is interesting that recent work has emphasized the importance of red-cell membrane flexibility in parasite invasion (Pasvol *et al.* 1988). By and large the membranes of the genetic variant red cells are relatively rigid. Clearly the relation between membrane flexibility and parasitic infection will be worth further exploration.

In a recent study carried out in Papua New Guinea, Oppenheimer *et al.* (1987) found that infants who had relatively high levels of haemoglobin Bart's at birth, and who therefore were assumed to be α^+ thalassaemia homozygotes, had significantly higher malarial parasite rates than normal infants or α^+ thalassaemia heterozygotes, both at 6 and 12 months. Spleen rates showed a similar trend at 6 months. On the other hand, homozygous α^+ thalassaemic infants showed a relatively higher haemoglobin level associated with malaria than the other groups. These are extremely interesting observations that need extending in a larger series. They suggest that, despite the higher parasitaemias, homozygotes for α^+ thalassaemia may be relatively protected. Whether, as the authors suggest, the higher parasite rates reflect a retarded rate of growth of parasites in the α thalassaemic red cells remains to be established. Alternatively these observations might indicate a more subtle immune protective effect. Another interesting observation from this study is that the effect of iron on increasing parasite rates was more marked in normal babies than in thalassaemics. Again, this raises the interesting question as to whether there is something about the small and relatively underhaemoglobinized thalassaemic red cell, perhaps its deformability, that provides a common mechanism for protection.

HLA-DR and infectious disease

Some of the published associations of HLA-DR with infectious disease are summarized in table 3. Here, we shall make no attempt to discuss the enormous literature on the well-defined HLA-DR associations with presumed autoimmune diseases such as ankylosing spondylitis and diabetes, conditions that may represent an abnormal immune response to environmental agents, including infection. Rather, we shall concentrate on the possibility of HLA-DR modification of specific infectious diseases.

There are several difficulties in interpreting some of the studies that are summarized in table 3. First, many of them involve extremely small numbers of patients; weak associations would have been missed and apparent associations may be spurious. And even where associations are stronger it is uncertain whether they reflect true HLA-DR relations or simply follow linkage disequilibrium between specific HLA-DR haplotypes and closely related genetic determinants on chromosome 6.

A surprising feature of HLA susceptibilty data on infectious disease is that few associations have been identified with class I (A, B and C locus) alleles. These alleles present viral antigens to cytotoxic T cells, which are the predominant defence against such infections. This lack of association may reflect a failure to study protection, rather than susceptibility to infection in

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

TABLE 3 HLA-DR ASSOCIATIONS WITH INFECTIOUS DISEASES OR AGENTS

(Ir indicates immune response, EB Epstein–Barr virus, and CAH chronic active hepatitis. References in Pollack & Rich (1985), Zabriskie & Gibofsky (1986), Ottenhoff *et al.* (1986), Abdel Salam *et al.* (1979), Jenkins & Dunn (1981), and C. Ludlam, personal communication. See text.)

disease	association	comment
hepatitis B infection	none	—
hepatitis B carrier	DR	?DR associated Ir
hepatitis B vaccine failure	DR7	?DR associated Ir
antigen positive CAH	decreased DR4	?DR associated Ir
infectious mononucleosis	none	—
EB with low VCA response	DR1	?DR associated Ir
AIDS; sero conversion in haemophilia	A1B8DR3	
AIDS + Kaposi sarcoma	DR5 or DR5/DR2	?DR associated Ir
T cell leukaemia	None	—
poliomyelitis	Bw16 increased	no HLA-D deviations DR not studied
tuberculosis	Bw35.B8DR2 B8 increased	excess haplotype sharing between affected siblings ?DR associated Ir
tuberculosis	DR4	based on skin test responsiveness
amoebic abscess	B16	—
P. falciparum malaria	A1, W5, W21	altitude-related HLA and haplotype relationship. DR not studied
schistosomiasis	A1, B5	associated with hepato splenomegaly
lyme disease	DR2	?DR associated Ir
lepromatous leprosy	DR1, DR2	DR associated Ir
tuberculous leprosy	DR2l, DR3	

large populations. Alternatively, it may indicate that serological specificities inadequately reflect the functional variants identified only when T cell responses to viral antigens are studied.

There are, however, features emerging from these studies that suggest that the HLA-DR system may be of considerable importance, certainly in terms of modifying the clinical course of common infectious or parasitic diseases (Pollack & Rich 1985; Zabriskie & Gibofsky 1986). For example, the persistent carrier status for hepatitis B virus, and the lack of response to vaccines against this agent, appear to be DR-associated, suggesting an unusual type of immune response in subsets of individuals in different populations (Rubenstein *et al.* 1982; Carbonara *et al.* 1983). There appears to be a strong relation between DR4 and antigen-positive chronic active hepatitis although unfortunately the published series have been small (Paronetto *et al.* 1982). The limited amount of published data on AIDS is particularly intriguing, suggesting as it does a DR association in the subgroup of patients who develop Kaposi's sarcoma. Similarly, the recent observations that, among a group of haemophiliacs who received factor VIII from one HIV-contaminated batch, there was a strong correlation between a particular HLA-DR haplotype and seroconversion suggests that genetic factors may be important in modifying the course of HIV infections (C. Ludlam, personal communication). Data on DR association and tuberculosis and Lyme disease are tantalizing, although again they depend on a relatively small number of observations in different populations (see Zabriskie & Gibofsky 1986).

Perhaps the most convincing example of modification of the pattern of a disease by products of the HLA-DR gene cluster are the observations made on patients with different forms of

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

leprosy (reviewed by de Vries *et al.* (1985) and Pollock & Rich (1985)). As this topic is discussed in detail elsewhere in this symposium (Fine, this symposium) we shall only summarize it briefly here. A number of independent reports have noted the association of an HLA-DR2-related determinant DQ1 with lepromatous leprosy and the association of DR3 with tuberculoid leprosy. And, by and large, family data studying segregation of cases of leprosy support the role of HLA factors in susceptibility to infection (see Pollock & Rich 1985). For example, the latter authors point out that in the family studies of van Eden *et al.* (1985) the data show a strong trend for non-random segregation of parental haplotypes among healthy children, as well as a highly significant non-random segregation among the children with leprosy. In fact, linkage analyses from two large studies reviewed by Pollock and Rich gives a combined lod score for linkage of HLA susceptibility to either form of leprosy of 3.10, with a recombination fraction, θ, of 0.30†.

How might the immune response-like gene modify the clinical course of leprosy? It turns out that antibody responses are high in lepromatous leprosy but low or absent in tuberculoid leprosy. Like other mycobacterial infections, host defence to *Mycobacterium leprae* is mediated mainly by T cells; patients with HLA-associated defective T cell responses develop lepromatous leprosy, characterized by high antibody response to *M. leprae* antigens, but specific skin test anergy. On the other hand, certain HLA alleles are associated with relatively more effective cell-mediated responses, presumably predisposing to the tuberculoid form of the disease. Overall, these differing immune responses suggest alternative consequences of T cell recognition of the antigens after definition in the context of self-HLA molecules (Pollock & Rich 1985).

What of other important tropical diseases? A neglected but important observation on the possible role of HLA in modifying susceptibility to malaria was published in 1972 by Piazza and colleagues. As mentioned earlier, the classical studies of Siniscalco and Ceppellini in Sardinia, in which they interpreted the difference in prevalance of thalassaemia and G6PD deficiency in populations from different altitudes as evidence of protection against malaria, remains one of the seminal studies in this field. In 1972 Ceppellini and his colleagues returned to Sardinia and resampled the villages that had comprised the earlier study (Piazza *et al.* 1972). In addition to confirming the difference in frequency between the red-cell polymorphisms, on this occasion they also showed that there was a significant difference in HLA types between the mountainous regions and the low-lying valleys. In summary, they found that there was a highly significant heterogeneity of gene frequency in HLA1, W5, and W21 between the highland and lowland populations. Furthermore, they found a striking difference in haplotype frequencies between these groups. Apart from the common red-cell polymorphisms there were no other major differences in polymorphisms between the populations. In fact, from their data the authors conclude that the highland and lowland settlements that were studied have a common and distinctive ancestry and that their genetic homogeneity may have been increased by selective forces. This, they conclude, adds weight to the hypothesis of the adaptive nature of the variations that discriminate between highland and lowland populations. To our knowledge this work has not been repeated elsewhere.

Other possible HLA associations with tropical diseases are summarized by Jenkins & Dunn (1981).

† A lod score represents the logarithm to the base 10 of the likelihood (L) ratio $L(\theta)/L(0.5)$.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Complement

Although there have been no studies of the distribution of complement polymorphisms in relation to infectious disease at the population level it is becoming apparent that inherited variability of the complement system, particularly the terminal steps in the pathway, may be associated with a marked susceptibility to recurrent infection, particularly with neisseria. The inherited complement abnormalities have been reviewed by Schur (1986) and will only be outlined here with respect to susceptibility to infection.

It is quite clear that deficiencies in the later parts of the complement pathway are associated with marked proneness to infection. Interestingly, these conditions are not uncommon. For example, over 24 families have been identified as having individuals with C6 deficiency, most of whom have been identified because of recurrent attacks of neisserial meningitis. Similar observations have been made in over 16 families with C7 or C8 deficiency. As well as these specific deficiencies it is becoming clear that each of the components of the complement system is polymorphic and a number of different electrophoretically determined allotypes have been identified. So far there is no evidence that they are related specifically to susceptibility to infection although there has been a description of the association of the F1 allele with lepromatous leprosy in certain mongoloid populations.

Clearly, the complement system offers an extremely interesting series of candidate genes for the further analysis of genetic heterogeneity in response to a wide variety of bacterial infections.

Blood groups

The enormous and conflicting literature on the relation between blood groups and infection is summarized ably by Mourant *et al.* (1978), Vogel & Motulsky (1982) and Mourant (1983).

One of the major problems in interpreting the complex distribution of the ABO blood groups in terms of selection is that it is impossible to envisage a biological situation that is nearly as simple as that described earlier for the haemoglobin variants and *P. falciparum* malaria. It seems likely that in many populations malaria has remained endemic for centuries, even several thousand years, because the ecological conditions for the mosquito vector have not changed. However, many bacterial and viral infections come and go as short explosive epidemics. Sometimes cataclysmic events will be long remembered, the plague epidemic in the Middle Ages for example. In other cases infections may not be recorded, for example the infantile diarrhoeas. Given the enormous diversity of infectious disease and these temporal problems, together with the inherent difficulties consequent on the many other ways in which blood group distributions may vary, it is not surprising that the epidemiological approach has been extremely difficult to apply to this problem. Apart from the homologies between certain of the blood group antigens and bacterial antigens there have been very few ways of testing susceptibility or resistance to bacterial or viral infection by specific blood groups directly.

A few extremely tentative conclusions have been drawn from a combination of population studies and more direct observations on smaller groups. These may be summarized as follows.

(i) The extremely high group O frequency in Central and South America may have been due to advantage of O in the presence of syphilis.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

(ii) The higher frequency of O in marginal European populations could be caused by lower selection against O carriers by plague.

(iii) The relatively low A frequencies in central and southern Asia might be due to selection by smallpox. In the same areas O is not so frequent, the gap being filled by B. This advantage of group B might be due to long-standing selection against A by smallpox as well as against type O by plague. Unfortunately, much of the evidence is conflicting and it is unlikely that it will be possible to improve such data in the future.

(iv) Secretor status may be an important factor in susceptibility to some bacterial infections (Blackwell *et al.* 1986).

There is even less evidence regarding the relation of blood groups to parasitic disease, with one notable exception. Miller *et al.* (1975) noted that human erythrocytes lacking the Duffy blood group (FyFy) are refractory to invasion by merozoites of *P. knowlesi*, a parasite with many similarities to the human malarial parasite, *P. vivax*. Unlike the rest of the world population West Africans are almost entirely Duffy negative. It turns out that Duffy-negative volunteers are completely refractory to experimental infection with *P. vivax*. Thus it seems likely that the paucity of the Duffy blood group in West Africa is the result of selection against *P. vivax*. Homozygotes are haematologically normal and therefore if they are resistant to this infection the Duffy-negative gene should have gone to fixation, as appears to be the case in West Africa. This situation is unlike the balanced polymorphisms described earlier for the other inherited red-cell disorders.

Interestingly, *P. vivax* malaria is now an extremely mild disease in West Africa. The fact that the Duffy-negative phenotype has reached fixation suggests that selection was probably extremely strong at some time in the past, a fact that emphasizes the occurrence of evolutionary changes at both the level of the red cell and of the parasite, and, incidentally, the problems of interpreting current poulation data in the light of past events. Pasvol & Wilson (1982) have summarized the evidence for and against the Duffy determinant being the receptor for *P. vivax*. It is, of course, possible that weaker selection over a long period of time could have resulted in the distribution of the Duffy group in Africa or, as suggested by Livingstone (1985), it could have occurred by drift without implicating strong selection by *P. vivax*.

Immunoglobulin genetic markers (allotypes)

There are two major unlinked genetic systems that are described under this category; Gm, controlling IgG-C-region heavy chain synthesis, and INV, controlling the C region of the κ chains, now called Km. In addition a heavy chain genetic marker, Am, has been found on IgA_2 subclass molecules. Very little is known about the distribution of the different subtypes of these systems in relation to infection; there appear to be some strong associations between antibody response to polysaccharide-encapsulated bacteria such as *Haemophilus influenzae*, pneumococcus and meningococcus, and certain Gm and Km variants.

Single gene disorders that predispose to infection

There is a wide variety of single-gene disorders that predispose to infection. Some of these are summarized in table 4. Most of them involve specific enzyme deficiencies in neutrophils or lymphocytes, or reflect monogenic defects in the synthesis of immunoglobulins.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

TABLE 4. SOME SINGLE GENE DISORDERS ASSOCIATED WITH INCREASED SUSCEPTIBILITY TO INFECTION

granulocyte/macrophage function	
adherence	– plasma membrane glycoprotein deficiency
chemotaxis	– complement deficiency C1–C5
degranulation	– Chediak–Higashi
	– specific granule deficiency
oxidative metabolism	– CGD (cytochrome b_6)
	– G6PD
	– myeloperoxidase
	– glutathione reductase
	– glutathione synthetase
	– glutathione peroxidase
primary B cell deficiency	
X-linked hypogammaglobulinaemia	
autosomal recessive hypogammaglobulinaemia	
selective hypogammaglobulinaemia IgG, IgA	
primary T cell deficiency	
severe combined	
adenosine deaminase deficiency	
purine nucleoside phosphorylase deficiency	
MHC class I or II deficiency	
with other major defects	
Wiskott–Aldrich. ataxia telangiectasia	
complement	
late pathway defects	

DISCUSSION

From this short review it is apparent that our understanding of the genetic factors that may modify susceptibility to infectious and parasitic disease is extremely flimsy. However, such models as have been studied in detail suggest that genetic susceptibility may be an important factor in modifying the response of populations to these conditions.

Overall, the malaria hypothesis has lasted the years remarkably well. It has been strengthened considerably by recent work utilizing recombinant DNA technology, and clearly there is scope for pursuing this approach in an attempt to analyse all the common red-cell polymorphisms. At least for the haemoglobin variants, thalassaemia and G6PD deficiency, it is clear that these conditions have had multiple origins in different populations and that malaria is likely to have played a major role in achieving their high gene frequencies.

P. falciparum malaria has been a major killer over the years. Thus even though the genetic red-cell variants have achieved polymorphic frequencies in many populations, the individual degree of protection afforded by them may be only quite small. This may partly explain why it has been so difficult to determine the cellular mechanisms involved. Even in the case of the sickle-cell gene we still have only the most rudimentary information about how such protection might be mediated. And although the development of *in vitro* culture systems has stimulated a great deal of work it has to be remembered that it is a long way from the *in vivo* situation. Furthermore, the observation that protection in the field can only be demonstrated during specific phases of development in African children raises another potential problem; by analysing the invasion and growth of parasites in adult cells in culture we may be missing the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

protective mechanism. Is there, for example, something unusual about red cells in foetal or neonatal life? One factor that may be important in this regard is that even in normal children the red cells are smaller during the early years of postnatal life than they are later on during development; this effect is magnified by the various thalassaemia genes. And of course it is also possible that some of the protective effect is mediated by more subtle immune mechanisms that will certainly be missed in any of these *in vitro* systems.

Because, as mentioned earlier, parasitic illnesses such as malaria have been endemic in populations over many thousands of years they offer a particularly valuable model system for analysing genetic variability in susceptibility to these agents. It is very unlikely that the red-cell polymorphic systems are unique; there must be similar selective factors at work in these populations. The HLA-DR system is a particularly strong candidate and the results of the studies on patients with leprosy, and the earlier work from Sardinia on malaria, suggest that it will be particularly fruitful to pursue studies relating polymorphisms of this system to parasitic disease. The application of RFLP analysis using probes for the class I and II genes of the HLA-DR gene cluster should offer a valuable adjuvant to serological studies. Although the mechanisms whereby variation in structure of the class I and II genes might modify response to infection are still not clear, the recent finding of sequences in human class I genes with a high degree of homology to human cytomegalovirus (Beck & Barrell 1988) offers a number of intriguing possibilities.

As compared with parasitic illnesses it may be much more difficult to analyse the genetic factors that may change susceptibility to bacterial or viral infections. Many of these conditions are characterized by epidemics that, although they may be catastrophic, are often short-lived. Although such an event may be reflected by the alteration of the genotype of part of a population by differing genetic susceptibility, it may be extremely difficult to identify such an occurrence after the event. On the other hand, during such an epidemic, particularly if there is a high morbidity or mortality such as is occurring with AIDS at the moment, it may be possible to identify subsets of individuals who are genetically prone or resistant to the infective organism. As we have seen from this review there is no shortage of candidate genes, many of which have now been cloned and for which RFLPs are available. It should be possible, therefore, to define at least some of the loci involved.

An alternative approach to assessing the importance of genetic factors in susceptibility to particular infections is to compare concordance rates in monzygotic and dizygotic twins. Here again, data are very limited but such early studies on both tuberculosis (Kallman & Reisner 1943) and poliomyelitis (Herndon & Jennings 1951) argue for the importance of genetic factors in these conditions. As more candidate gene probes become available the importance of each locus may be assessed by study of the segregation of DNA polymorphisms with disease susceptibility within families.

In addition to our rather limited evidence from human studies, work on mouse models indicates that a number of clearly defined loci are involved in modifying responses to infection. For example, inate susceptibility to *Leishmania donovani* infection, measured over 2–4 weeks, is under the control of a single autosomal gene (*Lsh*), segregating for incompletely dominant resistance (r) and recessive susceptibility (s) alleles. This locus segregates independently from the known histocompatibility loci to a position between the centromere and *Idh-1* on chromosome 1. It has been observed that homozygous recessive strains of mice show two patterns of recovery when the course of infection is followed over a longer time. In some strains

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

there is a dramatic fall in parasitic numbers associated with severe liver damage whereas in others high parasite loads which involve mononuclear phagocytes throughout the body are maintained for up to two years. It turns out that this difference in long-term response is largely controlled by a gene or genes within, or close to, the major histocompatibility complex (MHC) (H-2) in the mouse (Blackwell *et al.* 1980). Indeed, there is evidence that the MHC is involved in response to other parasites in the mouse, including intestinal worms. Specific loci involved with increased resistance to bacterial infection have also been defined in the mouse (see Skamene 1985). For example, there is a dominantly inherited form of resistance to salmonellosis that also maps on chromosome 1; in this case the locus is called *Ity* although the *Lsh* and *Ity* genes may well be identical (Blackwell 1985). A number of other loci have been defined that are involved in resistance or susceptibility to viral infections (Mouse Newsletter 1987).

It seems likely that similar specific loci for resistance or susceptibility to infectious or parasitic disease will be defined in man. The availability of large numbers of single-point RFLPs and probes for highly variable regions (HVRs) should enable this problem to be studied with some possibility of success.

Because of our extremely limited information about the relative importance of genetic factors in infection in man it is particularly difficult to predict their overall contribution in modifying the response to infection in large populations. It is even more difficult to define the way in which genetic variability might be involved in the transmission of infectious disease and hence in modifying the pattern of human epidemics. From the few models that are available in mouse and man it is clear that at least some forms of genetic resistance to infection or parasitic illness are associated with persistence of relatively high levels of organisms in the blood or tissues. Presumably this would make the particular pathogens more accessible to vectors and therefore might well have an important effect in modifying the pattern of transmission, both in endemic and epidemic conditions. These factors, taken together with the extraordinary facility for many organisms to change their own genetic makeup, underline the considerable difficulties in building adequate models of the spread of infectious disease; genetic factors, both in host and pathogen, must play an important role in modifying the course, both in individuals and in poulations.

References

Abdel Salam, E., Ishaac, S. & Mahmoud, A. A. F. 1979 Histocompatibility-linked susceptibility for hepatosplenomegaly in human schistosomiasis mansoni. *J. Immun.* **123**, 1829–1831.

Allison, A. C. 1954 Protection afforded by sickle cell trait against subtertian malarial infection. *Br. med. J.* **1**, 290–294.

Antonarakis, S. E., Boehm, C. D., Giardina, P. J. V. & Kazazian, H. H. 1982 Non random association of polymorphic restriction sites in the β-globin gene complex. *Proc. natn. Acad. Sci. U.S.A.* **79**, 137–141.

Antonarakis, S. E., Boehm, C. D., Serjeant, G. R., Theisen, C. E., Dover, G. J. & Kazazian, H. H. 1984 Origin of the β^s-globin gene in blacks: the contribution of recurrent mutation or gene conversion or both. *Proc. natn. Acad. Sci. U.S.A.* **81**, 853–856.

Beck, S. & Barrell, B. G. 1988 Human cytomegalovirus encodes a glycoprotein homolgous to MHC class-I antigens. *Nature, Lond.* **331**, 269–272.

Blackwell, C. C., Jonsdottir, K., Hanson, M., Todd, W. T. A., Chaudhuri, A. K. R., Mathew, B., Brettle, R. P. & Weir, D. M. 1986 Non-secretion of ABO antigens predisposing to infection by *Neisseria meningitidis* and *streptococcus pneumoniae*. *Lancet* ii, 284–285.

Blackwell, J. M. 1985 Genetic conversion of discrete phases of complex infections: *Leishmania donovani* as a model. In *Genetic control of host resistance to infection and malignancy* (ed. E. Skamene), pp. 31–49. New York: Alan R. Liss.

Blackwell, J., Freeman, J. & Bradley, D. 1980 Influence of H-2 complex on acquired resistance to *Leishmania donovani* infection in mice. *Nature, Lond.* **283**, 72–74.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Brockelman, C. R., Wongsattayanont, B., Tan-Ariya, P. & Fucharoen, S. 1987 Thalassemia erythrocytes inhibit *in vitro* growth of *Plasmodium falciparum*. *J. clin. Microbiol.* **25**, 56–60.

Brown, P. J. 1981 New considerations on the distribution of malaria, thalassemia, and glucose-6-phosphate dehydrogenase deficiency in Sardinia. *Hum. Biol.* **53**, 367–382.

Bunn, H. F. & Forget, B. G. 1986 *Hemoglobin: molecular, genetic and clinical aspects*. Philadelphia: W. B. Saunders.

Carbonara, A., Mayr, W., Rizzetto, M., Contu, L., Curtoni, E. S., de Marchi, M., Farci, P., Lavarini, C. & Olivetti, E. 1983 Endemic HBV infection, tissue autoantibodies and HLA. Analysis of a Sardinian population. *Tissue Antigens* **22**, 289–294.

de Vries, R. R. P., van Eden, W. & Ottenhoff, T. H. M. 1985 HLA class-II immune response genes and products in leprosy. *Prog. Allergy* **36**, 95–113.

Fleming, A. F., Storey, J., Molineaux, L., Iroko, E. A. & Attai, E. D. E. 1979 Abnormal haemoglobins in the Sudan savanna of Nigeria. I. Prevalence of haemoglobins and relationships between sickle cell trait, malaria and survival. *Ann. trop. Med. Parasitol.* **73**, 161–172.

Flint, J., Hill, A. V. S., Bowden, D. K., Oppenheimer, S. J., Sill, P. R., Serjeantson, S. W., Bana-Koiri, J., Bhatia, K., Alpers, M. P., Boyce, A. J., Weatherall, D. J. & Clegg, J. B. 1986 High frequencies of α thalassaemia are the result of natural selection by malaria. *Nature, Lond.* **321**, 744–749.

Friedman, M. J. 1978 Erythrocytic mechanism of sickle cell resistance to malaria. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1994–1997.

Friedman, M. J. 1979 Oxidant damage mediates variant red cell resistance to malaria. *Nature, Lond.* **280**, 245–249.

Friedman, M. J., Roth, E. F., Nagel, R. L. & Träger, W. 1979 *Plasmodium falciparum*: physiological interactions with the human sickle cell. *Expl Parasitol.* **47**, 73–80.

Friedman, M. J. & Träger, W. 1981 The biochemistry of resistance to malaria. *Scient. Am.* **244** (2), 154–155.

Haldane, J. B. S. 1949 In *Symposium on ecological and genetic factors in animal speciation. Pallanza, 1948*, pp. 68–76. Supplemento a *La rocerca scientifica* **19**.

Herndon, C. N. & Jennings, R. G. 1951 A twin family study of susceptibility to poliomyelitis. *Am. J. hum. Genet.* **3**, 17–46.

Higgs, D. R., Hill, A. V. S., Bowden, D. K., Weatherall, D. J. & Clegg, J. B. 1984 Independent recombination events between duplicated human α globin genes: implications for their concerted evolution. *Nucl. Acids Res.* **12**, 6965–6967.

Higgs, D. R. & Weatherall, D. J. 1983 Alpha thalassemia. *Curr. Top. Hematol.* **4**, 37–97.

Hill, A. V. S. 1986 The population genetics of α thalassemia and the malaria hypothesis. *Symposium on quantitative biology* (ed. J. D. Watson), pp. 489–498. Cold Spring Harbor.

Hill, A. V. S., Bowden, D. K., O'Shaughnessy, D. F., Weatherall, D. J. & Clegg, J. B. 1988 β Thalassemia in Melanesia: association with malaria and characterization of a common variant (IVS1 *nt* 5 G→C). *Blood*. (In the press.)

Hill, A. V. S., Bowden, D. K., Trent, R. J., Higgs, D. R., Oppenheimer, S. J., Thein, S. L., Mickelson, K. N. P., Weatherall, D. J. & Clegg, J. B. 1985 Melanesians and Polynesians share a unique α thalassemia mutation. *Am. J. hum. Genet.* **37**, 571–580.

Jenkins T. & Dunn, D. S. 1981 Haematological genetics in the tropics. Tropical Africa. *Clin. Haematol.* **10**, 1029–1050.

Kan, Y. W. & Dozy, A. M. 1980 Evolution of the hemoglobin S and C genes in world populations. *Science, Wash.* **209**, 388–391.

Kallman, F. J. & Reisner, D. 1943 Twin studies on the significance of genetic factors in tuberculosis. *Ann. Rev. Tuberc.* **47**, 549–574.

Kazazian, H. H., Waber, P. G., Boehm, C. D., Lee, J. I., Antonarakis, S. E. & Fairbank, V. F. 1984 Hemoglobin E in Europeans – further evidence for multiple origins of the beta-E-globin gene. *Am. J. hum. Genet.* **36**, 212–217.

Kidson, C., Lamont, G., Saul, A. & Nurse, G. T. 1981 Ovalocytic erythrocytes from Melanesians are resistant to invasion by malarial parasites in culture. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5829–5832.

Kulozik, A. E., Kar, B. C., Satapathy, R. K., Serjeant, B. E., Serjeant, G. R. & Weatherall, D. J. 1987 Fetal hemoglobin levels and β^s globin haplotypes in an Indian population with sickle cell disease. *Blood* **69**, 1742–1746.

Kulozik, A. E., Wainscot, J. S., Serjeant, G. R., Kar, B. C., Al-Awamy, B., Essan, G. J. F., Falusi, A. G., Haque, S. K., Hilali, A. M., Kate, S., Ranasinghe, W. A. E. P. & Weatherall, D. J. 1986 Geographical survey of β^s-globin gene haplotypes: evidence for an independent Asian origin of the sickle-cell mutation. *Am. J. hum. Genet.* **39**, 239–244.

Livingstone, F. B. 1985 *Frequencies of hemoglobin variants*. New York: Oxford University Press.

Luzzatto, L. 1985 Malaria and the red cell. In *Recent advances in haematology* 4th edn (ed. A. V. Hoffbrand), pp. 109–126. Edinburgh: Churchill Livingstone.

Luzzato, L., Nwachuku, E. S. & Reddy, S. 1970 Increased sickling of parasitized erythrocytes is mechanism of resistance against malaria in the sickle trait. *Lancet* i, 319–321.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Luzzatto, L., Sodiende, O. & Martini, G. 1983 Genetic variation in the host and adaptive phenomena in *Plasmodium falciparum* infection. In *Malaria and the red cell* (*CIBA Foundation Symposium*), pp. 159–173. Amsterdam: Elsevier/North Holland.

Mears, J. G. & Lachman, H. M. 1981 Sickle gene: its origin and diffusion from West Africa. *J. clin. Invest.* **68**, 606–610.

Miller, J., Golenser, J., Spira, D. T. & Kosower, N. S. 1984 *Plasmodium falciparum*: thiol status and growth in normal and glucose-6-phosphate dehydrogenase deficient human erythrocytes. *Expl Parasitol.* **57**, 239–247.

Miller, L. H., Mason, S. J., Dvorak, J. A., McGinniss, M. H. & Rothman, I. K. 1975 Erythrocyte receptors for (*Plasmodium knowlesi*) malaria: Duffy blood group determinants. *Science, Wash.* **189**, 561–563.

Mourant, A. E. 1983 *Blood relations: blood groups and anthropology*. Oxford University Press.

Mourant, A. E., Kopec, A. C. & Domaniewska-Sobczak, K. 1978 *Blood groups and diseases*. Oxford University Press.

Mouse News Letter 1987 No. 77. Oxford University Press.

Oppenheimer, S. J., Hill, A. V. S., Gibson, F. D., MacFarlane, S. B., Moody, J. B. & Pringle, J. 1987 The interaction of alpha thalassaemia with malaria. *Trans. R. Soc. trop. Med. Hyg.* **81**, 322–326.

Oppenheimer, S. J., Higgs, D. R., Weatherall, D. J., Barker, J. & Spark, R. A. 1984 α Thalassaemia in Papua New Guinea. *Lancet* i, 424–426.

Orkin, S. H. & Kazazian, H. H. 1984 The mutation and polymorphism of the human β-globin gene and its surrounding DNA. *Ann. Rev. Genet.* **18**, 131–171.

Ottenhoff, T. H. M., Torres, P., de Las Agnas, J. T., Fernandez, R., van Eden, W., de Vries, R. R. P. & Stanford, J. L. 1986 Evidence for an HLA-DR4-associated immune-response-gene for *Mycobacterium tuberculosis*. *Lancet* ii, 310–312.

Paronetto, F., Piccinino, F., Vernace, S., Aloisio, V., Giusti, G. & Rubenstein, P. 1982 HLA antigens in HBsAG-positive chronic active hepatitis (CAH): population and family studies. In *Viral hepatitis 1981 International Symposium*, pp. 728–729. Philadelphia: Franklin Inst. Press.

Pasvol, G., Chasis, J. A., Mohandas, N., Anstee, D. J., Tanner, M. J. A., Mawby, W. J. & Merry, A. H. 1988 Monoclonal antibodies and Fab fragments to the external portion of glycophorin may inhibit the invasion of red cells by *Plasmodium falciparum* by decreasing membrane deformability. (Submitted.)

Pasvol, G., Weatherall, D. J. & Wilson, R. J. M. 1977 Effects of foetal haemoglobin on susceptibility of red cells to *Plasmodium falciparum*. *Nature, Lond.* **270**, 171–173.

Pasvol, G. & Weatherall, D. J. 1978 A mechanism for the protective effect of haemoglobin S against *P. falciparum* malaria. *Nature, Lond.* **274**, 701–703.

Pasvol, G. & Wilson, R. J. M. 1982 The interaction of malaria parasites with red blood cells. *Br. med. Bull.* **38**, 133–140.

Piazza, A., Belvedere, M. C., Bernoco, D., Conighi, C., Contu, L., Curtoni, E. S., Mattiuz, P. L., Mayr, W., Richiardi, P., Scudeller, G. & Ceppellini, R. 1972 HL-A variation in four Sardinian villages under differential selective pressure by malaria. In *Histocompatibility testing*, pp. 73–84. Copenhagen: Munksgaard.

Pollack, M. S. & Rich, R. R. 1985 The HLA complex and the pathogenesis of infectious diseases. *J. infect. Dis.* **151**, 1–8.

Pressley, L., Higgs, D. R., Clegg, J. B. & Weatherall, D. J. 1980 Gene deletions in α thalassemia prove that the 5′ ζ locus is functional. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3586–3589.

Roth, E. F. Jr. Raventos-Suarez, C., Rinaldi, A. & Nagel, R. L. 1983 Glucose-6-phosphate dehydrogenase deficiency inhibits *in vitro* growth of *Plasmodium falciparum*. *Proc. natn. Acad. Sci. U.S.A.* **80**, 298–299.

Rubenstein, P., Walker, M., Krassner, J., Carrier, C., Carpenter, C., Stevens, C., Szmuness, W. & Baker, L. 1982 HLA and hepatitis-B (HB) responsiveness. Comparative studies in vaccine recipients and in persistent antigen and antibody carriers. *Hum. Immunol.* **5**, 168.

Skamene, E. 1985 *Genetic control of host resistance to infection and malignancy*. New York: Alan R. Liss.

Schur, P. H. 1986 Inherited complement component abnormalities. *Ann. Rev. Med.* **37**, 333–346.

Siniscalco, M., Bernini, L., Filippi, G., Latte, B., Khan, M., Piomelli, S. & Rattazzi, M. 1966 Population genetics of haemoglobin variants, thalassemia and glucose-6-phosphate dehydrogenase deficiency, with particular reference to malaria hypothesis. *Bull. Wld Hlth Org.* **34**, 379–393.

van Eden, W., Gonzalez, N. M., de Vries, R. R. P., Convit, J. & van Rood, J. J. 1985 HLA-linked control of predisposition to lepromatous leprosy. *J. infect. Dis.* **151**, 9–14.

Vogel, F. & Motulsky, A. G. 1982 *Human genetics: problems and approaches*. Heidelberg: Springer-Verlag.

WHO Working Group 1982 Hereditary anaemias: genetic basis, clinical features, diagnosis and treatment, *Bull. Wld Hlth Org.* **60**, 643–659.

Wainscoat, J. S., Bell, J. I., Thein, S. L., Higgs, D. R., Serjeant, G. R., Peto, T. E. A. & Weatherall, D. J. 1983 Multiple origins of the sickle mutation: evidence from β^s globin gene cluster polymorphisms. *Molec. Biol. Med.* **1**, 191–197.

Wainscoat, J. S., Hill, A. V. S., Boyce, A., Flint, J., Hernandez, M., Thein, S. L., Old, J. M., Lynch, J. R., Falusi, Y., Weatherall, D. J. & Clegg, J. B. 1986 Evolutionary relationships of human populations from an analysis of nuclear DNA polymorphisms. *Nature, Lond.* **319**, 491–493.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Weatherall, D. J. & Clegg, J. B. 1981 *The thalassaemia syndromes*, 3rd edn. Oxford: Blackwell Scientific Publications.

Weatherall, D. J., Clegg, J. B., Higgs, D. R. & Wood, W. G. 1988 The hemoglobinopathies. In *Metabolic basis of inherited disease* (ed. C. R. Scriver, A. L. Beaudet, W. S. Sly & S. Valle). New York: McGraw Hill Book Co. (In the press.)

Zabriskie, J. B. & Gibofsky, A. 1986 Genetic control of the susceptibility to infection with pathogenic bacteria. *Curr. Top. Microbiol. Immunol.* **124**, 1–20.

Discussion

C. C. Blackwell (*Department of Bacteriology, University of Edinburgh Medical School, U.K.*). Sir David referred to our findings of an association between non-secretion of the ABO blood group antigens and susceptibility to bacterial meningitis, recurrent urinary tract infections and superficial fungal infections. In addition, we have found a significant increase in the proportion of non-secretors among patients with several autoimmune diseases for which infectious aetiologies have been postulated: ankylosing spondylitis, psoriatic anthropathy, insulin-dependent diabetes and Graves disease. As each of these also has particular HLA associations, we have suggested that some common compromise(s) in the innate defences of non-secretors might result in the initial susceptibility to the infectious agent(s), and the HLA-governed response to the agent contributes to the pathogenesis of the diseases. Identification of non-secretion as a susceptibility factor has provided the basis for a number of hypotheses we are currently testing.

(i) The carbohydrate moieties of the ABO or Lewis^{b} antigens in body fluids of secretors might bind to lectin-like adhesins on microorganisms and reduce colonization of epithelial surfaces.

(ii) The Lewis^{a} antigen found predominantly on the cells of non-secretors might act as a receptor for some microorganisms.

(iii) As the structural component for the third complement component (C3) is in the same linkage group as the secretor gene, there might be quantitative or functional differences in the C3 of secretors and non-secretors.

(iv) Increases in the levels of total serum IgA or IgG observed among non-secretor women with recurrent urinary tract infections suggest there might be differences in quantity, class or subclass of the antibody responses of secretors and non-secretors to colonization or infection of mucosal surfaces.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES