GENETICS OF SUSCEPTIBILITY TO HUMAN INFECTIOUS DISEASE

Graham S. Cooke and Adrian V. S. Hill

Before Robert Koch's work in the late nineteenth century, diseases such as tuberculosis and leprosy were widely believed to be inherited disorders. Heritability of susceptibility to several infectious diseases has been confirmed by studies in the twentieth century. Infectious diseases, old and new, continue to be an important cause of mortality worldwide. A greater understanding of disease processes is needed if more effective therapies and more useful vaccines are to be produced. As part of this effort, developments in genetics have allowed a more systematic study of the impact that the human genome and infectious disease have on each other.

The observation that some infectious diseases seem to have an inherited element is not new, nor is the observation that individuals respond differently to particular infections. The behaviour of infectious agents can vary so greatly between strains that the effects of individual variation are best seen when the same strain of an organism simultaneously infects previously unexposed individuals. This has sometimes occurred after deliberate inoculation and sometimes through unforeseen error. Early experience of therapeutic infection with malaria to treat syphilis led to the observation that individuals differed in their response to the *Plasmodium* parasite that causes malaria[1]. Similar differences can now be seen in volunteers infected during malarial vaccine trials. In 1927, the accidental administration of *Mycobacterium tuberculosis* to a population in Lubeck, Germany, left some individuals unaffected, whereas it led to severe disease and death in others. During the 1980s, haemophiliacs were unknowingly given human immunodeficiency virus (HIV)-infected blood products and it soon became clear that the rate at which they progressed to AIDS differed between individuals.

Differences in susceptibility to disease can also be observed at the level of populations. The Fulani, a West African tribe, seem markedly more resistant to severe malaria than do neighbouring ethnic groups[2]. In addition, after a tuberculosis outbreak in an American nursing home, significant differences in infection rates were seen in different ethnic groups in the same environment[3].

Understanding the molecular basis for these differences should provide useful insight into infectious disease pathogenesis and aid the continuing global struggle to control these diseases. In this review, we focus on common susceptibility variants, and also discuss some single-gene disorders that are associated with altered susceptibility to infectious disease. Both provide important insights into natural immunity. Also reviewed are specific genes that have been implicated in the outcome of infectious diseases, and the genome-wide approaches that are being used to discover new candidate genes. (TABLE 1 lists the infectious and parasitic agents discussed in this article.)

Measuring genetic effects

Adoption and twin studies, traditional tools for examining the genetic epidemiology of complex diseases, have been effectively applied to infectious diseases. A Danish study examined the causes of death among adopted children and compared them with the causes of death in their biological and adoptive parents[4]. The early death of a biological parent from infection increased the risk of death of the child from an infectious disease nearly sixfold, consistent with a substantial genetic effect. Twin studies have compared the disease status among identical and non-identical twins,

Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK. Correspondence to A.V.S.H e-mails: grahamc@well.ox.ac.uk; adrian.hill@molecular-medicine.oxford.ac.uk

Table 1 | **Parasitic and infectious organisms and their effects**

Group of organisms	Organism	Comments
Protozoa	*Plasmodium falciparum*	A mosquito-borne parasite that is responsible for most of the severe forms of malaria and is a major global cause of death.
	Plasmodium vivax	A mosquito-borne parasite that causes a benign, rarely fatal form of malaria.
	Cryptosporidium parvum	Protozoan that causes bowel infection, usually in the setting of immunosuppression.
Fungi	*Pneumocystis carinii*	Causes lung infection, usually in the setting of immunosuppression.
Helminths	*Schistosoma mansoni*	Infection through contact with freshwater that contains larvae can lead to schistosomiasis (also known as bilharzia).
Mycobacteria	*Mycobacterium tuberculosis*	Organism that causes tuberculosis (TB), a major cause of death worldwide.
	Mycobacterium leprae	Organism that causes leprosy in one of its two forms — localized (tuberculoid) or disseminated (lepromatous).
	BCG	A strain of mycobacterium that is derived from *M. bovis* and used as a vaccine in the prevention of TB and leprosy.
Bacteria	*Vibrio cholerae*	A water-borne bacterium that can colonize the bowel and produce toxin, which gives rise to cholera.
	Neisseria meningitidis	Organism that is responsible for major outbreaks of bacterial meningitis worldwide.
	Salmonella typhi	Enteric bacterium that, following ingestion, can lead to typhoid.
	Yersinia pestis	A bacterium that is responsible for outbreaks of plague.
Viruses	Hepatitis B/C	Viruses that can lead to both acute hepatitis or chronic liver disease, including cirrhosis and cancer.
	HIV	Retrovirus that leads to the development of AIDS.

BCG, Bacille Calmette–Guerin; HIV, human immunodeficiency virus.

with the expectation that disease concordance will be higher in identical twins for a disease with a genetically determined component. This prediction has been borne out for some diseases, and HERITABILITY has generally been easier to show in chronic disorders with low infectivity (TABLE 2).

SIBLING RISK (λs) can be estimated for infectious diseases. In comparison to diseases such as inflammatory bowel disease and diabetes, the estimates of λs for infectious diseases seem generally quite modest. With infectious, and some non-infectious, diseases the increased risk produced by a shared environment is included in these measurements. Nonetheless, marked progress has been made in the identification of important genes[5,6].

So far, the existence of so-called 'major' susceptibility genes that account for a significant proportion of the genetic contribution to disease susceptibility at a population level remains to be shown. Segregation analyses of large, affected pedigrees for diseases such as leprosy, tuberculosis, hepatitis B and SCHISTOSOMIASIS indicate that such genes might exist, but it is not clear whether these results can be extrapolated to the population as a whole. Moreover, none of these genes has been identified despite continuing research. Such segregation analyses are not without their limitations[7], and many of the observed data are just as well explained by numerous minor genes or 'polygenes' — a finding that might tally better with our understanding of infectious disease as an important selective force for evolution.

Infectious disease and evolution

Several common genetic disorders have been associated with protection from infectious disease, suggesting that their continued presence in the population has been the result of selection by infectious agents. Not all these genes are expressed in tissues that would normally be considered to be part of the immune system and, indeed, polymorphisms in genes that are expressed in many tissues might be the result of selection by infectious agents.

For a disease to exert selective pressure it would have to have a significant effect on morbidity and mortality before reproductive age and to have been exerting these effects for long periods of time. Malaria and the effects

Table 2 | **Results from selected twin studies of infectious disease**

| Disease | Country | Concordance* | | Reference |
		MZ (%)	DZ (%)	
Tuberculosis	Germany	65	25	115
	USA	62	18	116
	UK	32	14	117
Leprosy	India	52	22	103
Poliomyelitis	USA	36	6	118
Hepatitis B	Taiwan	35	4	119

*Concordance refers to the probability that one twin develops disease if the other is affected. Identical twins are monozygotic (MZ) and non-identical twins are dizygotic (DZ).
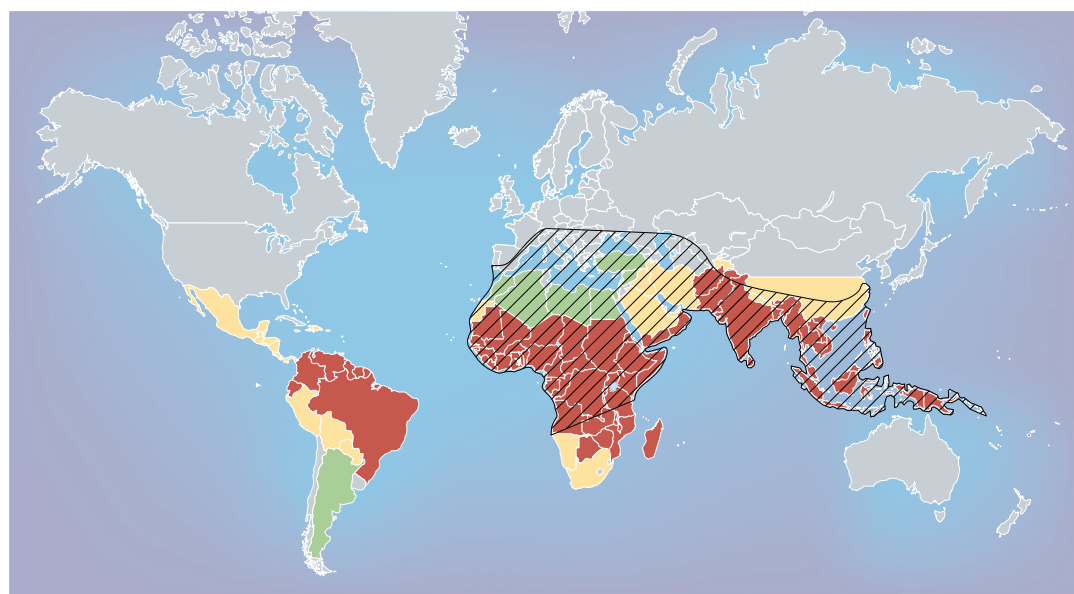
Figure 1 | **Global distribution of malaria and red-blood-cell disorders.** Green indicates areas where malaria is only present in a few remote locations, yellow indicates areas with intermediate malaria risk and red indicates areas with high malaria risk. The hatched area shows the distribution of red-blood-cell disorders. Source of data: World66.com and REF. 13.

of its causative agent, *Plasmodium falciparum*, on red-blood-cell diseases (haemoglobinopathies) are the best example. Disorders of red blood cells have been widely studied as they contribute a huge public health burden, particularly in Africa and Asia. They were among the earliest genetic diseases to be characterized molecularly, owing to their visible clinical and laboratory phenotypes. Early observational studies noted the similarity in geographical distribution of haemoglobinopathies and *P. falciparum* infection (FIG. 1). From this followed the hypothesis of J. B. S. Haldane[8,9], in which he proposed that red-blood-cell disorders, such as thalassaemia, might protect an individual from life-threatening infection with malaria.

Individuals that are homozygous for the HbS variant of haemoglobin (sickle haemoglobin) suffer the consequences of sickle-cell disease, but heterozygosity at this locus is strongly protective against severe malaria. Early studies that show the protective effect of HbS against *P. falciparum*[10] were the first to provide experimental evidence for a selective advantage to any genetic disorder. In West Africa, the greatest impact of HbS seems to be to protect against death and severe disease, with less effect on infection *per se*[11]. Its protective effects are probably due to impaired entry into, and growth of parasites in, affected erythrocytes[12].

Disorders of globin synthesis are found commonly in malarious regions and the disorders that give rise to the milder phenotypes are among the most common, single-gene disorders worldwide[13]. Population studies suggest that both α- and β-thalassaemias protect against malaria[14,15] but, in contrast to HbS, the mechanism of protection is less clear. That the study of such disorders can still illuminate the complex mechanisms of immunity to malaria is illustrated by a study in Vanuatu, in which an increased susceptibility to mild *Plasmodium vivax* and *P. falciparum* infections was observed among young children with α+-thalassaemia[16]. The authors have proposed that more frequent infections of immature red blood cells in young children might protect them against later life-threatening illness.

Glucose 6-phosphate dehydrogenase (G6PD) deficiency — an X-linked enzymatic disorder — hinders the capacity of red blood cells to deal with oxidative stress. Many sufferers experience HAEMOLYTIC crisis in response to certain foods or medicines. Many different mutations exist that result in reduced enzyme activity. Males that are hemizygous and females that are heterozygous for such mutations were found to be protected against severe malaria in both East and West Africa[17]. This benefit might result from impaired parasite growth in the erythrocytes[18] or from more efficient phagocytosis of parasitized red blood cells at an early stage of parasite maturation (FIG. 2).

Blood groups, both common and rare, have been associated with infectious disease. Blood group O (of the ABO system) is found at a higher frequency in severe cases of cholera than in the general population[19]. Although initial suggestions that this was related to the secretion of blood-group antigens into mucosal fluids were not confirmed, so-called SECRETOR STATUS probably does have a role in defence against some infections.

The Duffy blood group is a genetic variant that disrupts the Duffy antigen/chemokine receptor (DARC, also known as FY) promoter and alters a GATA1-binding site, which inhibits DARC expression on erythrocytes[20] and, therefore, prevents DARC-mediated entry of *P. vivax* into erythrocytes[21]. Although very common in West Africa, infection by this more benign malarial species does not typically

HAEMOLYSIS
Breakdown of red blood cells.

SECRETOR STATUS
Individuals are secretors or non-secretors depending on whether blood-group antigens are secreted into mucosal fluids.
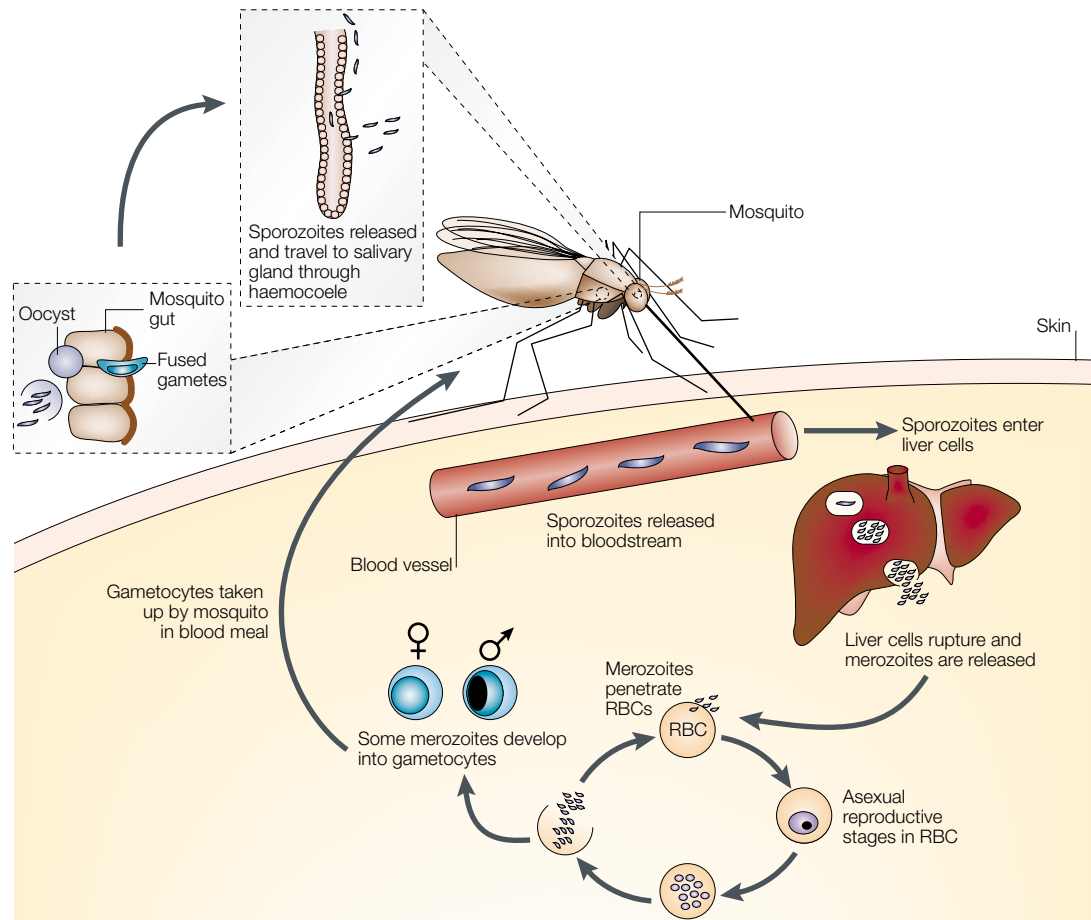
Figure 2 | **Life cycle of the malarial parasite *Plasmodium falciparum*.** *Plasmodium* spends one part of its life cycle in the human host, where it spreads through the circulation and enters the liver. In a subsequent step, liver cells rupture and merozoites are released into the blood stream. Within the red blood cells (RBCs), some merozoites develop into gametes that can be taken up by a mosquito in a blood meal. The sexual part of the life cycle is completed in the insect vector, where the gametes fuse and oocysts develop in the gut of the insect. Sporozoites are released and travel to the salivary glad, from where they can be transferred into the human host. Modified with permission from REF. 113 © (2001) Macmillan Magazines Ltd.

POPULATION BOTTLENECK
A marked reduction in population size followed by the survival and expansion of a small random sample of the original population.

MACROPHAGE
Phagocytic cell of the mononuclear lineage that internalizes and destroys infectious agents. Macrophages also function in antigen presentation.

CD40 LIGAND
Member of the tumour-necrosis factor superfamily of molecules that are expressed on the surface of T cells. CD40–CD40 ligand interaction is crucial for the development of many aspects of the immune system.

HYPER-IGM
The presence of unusually high levels of immunoglobulin M (IgM) in the blood.

cause severe disease or death. Unless it has done so at some stage in the past, there might be another explanation for the high prevalence of individuals in Africa who do not carry the Duffy antigen.

There are several common recessive disorders, such as cystic fibrosis, for which no evolutionary advantage for heterozygous carriers has been established. Selective pressure that maintains disease alleles need not necessarily be from infection, nor does it need to be acting at all. In theory, the high frequency of cystic fibrosis in Western Europe might be the result of a high mutation rate or a POPULATION BOTTLENECK, but the prevalence of diverse mutations is probably best explained by a survival advantage for carriers of mutant alleles[22,23]. How this survival advantage is achieved is not known, but there is accumulating evidence that the cystic fibrosis transmembrane conductance regulator (CFTR) has a role in the internalization of both *Pseudomonas aeruginosa* in the lung[24] and *Salmonella typhi* in the gut epithelium[25] (TABLE 1). Therefore, the high incidence of cystic fibrosis might be an evolutionary consequence of some protection from typhoid or another gut infection, such as

cholera. This factor, along with the varied demographic histories of European populations, probably accounts for the geographical distribution of the mutant alleles.

Speculation on evolutionary selection by infectious diseases is not limited to single-gene disorders. Mutations at major susceptibility loci for complex diseases might also have been influenced through selective pressure by infectious disease. *NOD2* — a gene recently implicated in Crohn disease[26,27] — is one example. It is predominantly expressed in MACROPHAGES, and it is possible that the florid inflammation seen in Crohn disease might, in the setting of particular infections, be advantageous.

Susceptibility and Mendelian disease
Among the wide variety of clinical disorders that are characterized by generalized immunodeficiency, modern genetic approaches have identified rare Mendelian conditions that predispose an individual to particular infectious diseases[28]. For example, CD40 LIGAND mutations have been identified as the basis for X-linked immunodeficiency with HYPER-IGM[29,30]. Patients with this condition seem susceptible to *Pneumocystis carinii* and

Box 1 | **The major histocompatibility complex**

The function of the human leukocyte antigens (HLAs), which are encoded by the major histocompatibility complex, is to bind peptide fragments that are produced as a result of intracellular protein degradation. HLA then displays these fragments on the cell surface where peptides are recognized by the T cells of the host. The HLA lies on chromosome 6 and extends over 4 Mb of DNA.
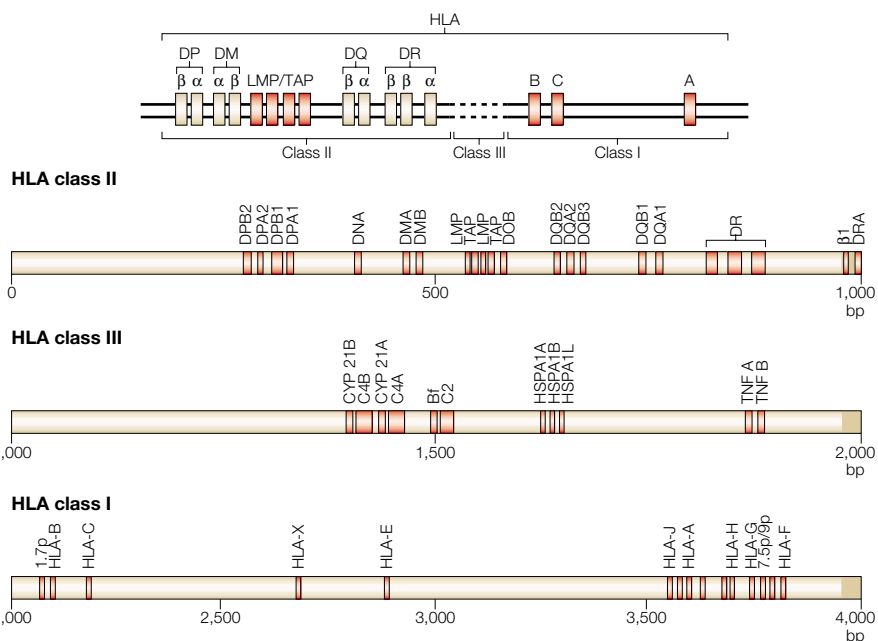
A, B and C are subclasses in the HLA class I region, and DR, DQ, DM and DP are subclasses of the HLA class II. Genes in the class III region encode complement proteins (C4A, C4B, C2 and Bf), tumour-necrosis factor-α (TNF-α) and TNF-β, enzymes involved in steroid synthesis (CYP21A and CYP21B) and heat-shock proteins (HSPA1A, HSPA1B and HSPA1L). Different HLA alleles have been associated with susceptibility or resistance to certain infections. Specific class I alleles have been associated with human immunodeficiency virus progression; B27 and B57 are associated with a good prognosis, whereas B35 is associated with a poorer one. The DRB1*1302 class II allele has been associated with spontaneous clearance of the hepatitis B virus. LMP, large multifunctional proteases; TAP, transporters associated with antigen processing. Modified with permission from REF. 112.

Cryptosporidium parvum infection (organisms that a healthy immune system will normally resist). Other families have provided important insights into mycobacterial immunity: one family has been identified that carries a recessive mutation in the interferon-γ receptor 1 (IFNγR1), which leaves homozygotes susceptible to severe infection by normally innocuous environmental mycobacteria[31]. Families with deletions in the *IFNGR1* gene show dominant susceptibility to OPPORTUNISTIC non-tuberculous mycobacterial and salmonella infections[32]. A similar phenotype is seen in individuals with IL12 (interleukin-12) deficiency[33] or with a mutation in the IL12 receptor[34]. More recently, a family has been identified in which a single locus seems to govern dominantly inherited susceptibility to *M. tuberculosis*[35].

Such families offer interesting insights into the molecules that are crucial for immunity to intracellular organisms. It is not known whether such highly penetrant mutations account for differences in susceptibility to disease in the general population, but it is unlikely that major susceptibility alleles would survive selective pressure from such ubiquitous organisms. More common alleles of these susceptibility loci, with a more subtle effect on gene function, are plausible sources of the variation that underlies differences between individuals in susceptibility to major pathogens. However, the evidence

so far does not support this possibility and the characterization of genes that underlie rare monogenic susceptibility to mycobacterial infection illustrates this point.

Susceptibility effects can be quite specific for particular microbial species. Individuals that are homozygous for non-functional *IFNGR1* alleles are at a markedly increased risk of infection by non-tuberculous mycobacteria. However, it is unclear whether they are at any increased risk of tuberculosis. So far, no significant linkage or associations between polymorphisms in the IFNGR1 genes and tuberculosis or leprosy have been identified. Possible biological reasons for this specificity are being uncovered. For example, IFNγ affects the growth of BCG (Bacille Calmette–Guerin, an attenuated bovine mycobacterium) but not virulent *M. tuberculosis* in human macrophages, and Ernst and co-workers[36] have identified a specific effect of *M. tuberculosis* action on the STAT1-signalling pathway that might underlie differential susceptibility to mycobacterial species. Further evidence for differing effects of STAT1 function between microbial species has come from the recent observation that *STAT1* mutations can affect mycobacterial but not viral immunity[37].

Human leukocyte antigen genes
Genes that are involved in protective immunity seem to have been under greater selective pressure, and show

OPPORTUNISTIC INFECTION
An infection that is normally resisted by a healthy individual but takes hold in the setting of a compromised immune system.

STAT FACTOR
(signal transducer and activator of transcription). Molecule that comprises one part of an intracellular pathway that mediates the effects of interferon-γ and other cytokines.
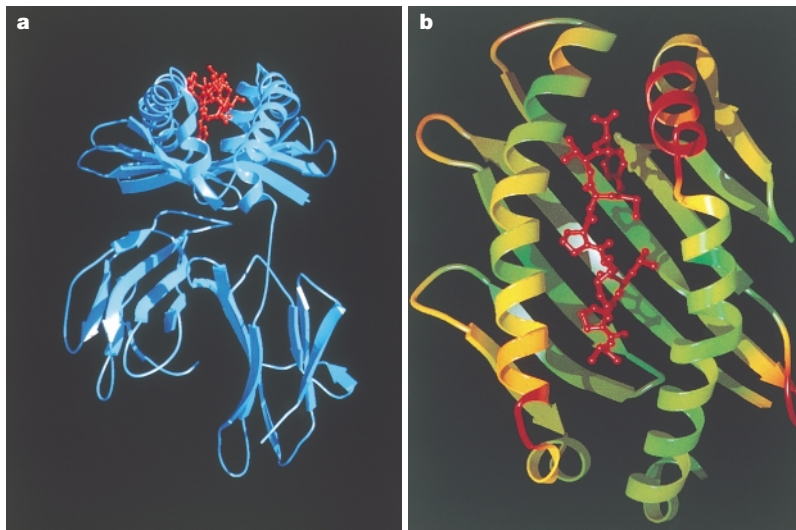
Figure 3 | **Structural diagram of HLA B53 (class I) molecule. a** | Here, the molecule is shown binding a peptide (red) from *Plasmodium falciparum*. **b** | Alternative view from above the HLA-binding groove. HLA, human leukocyte antigen. Reproduced with permission from REF. 114 © (1996) Cell Press.

greater variability, than other genes[38]. It is probable that many common biochemical variations have been selected by pressure from infectious diseases. In the search for the specific genetic polymorphisms that influence infectious disease, many candidate genes identified on the basis of their known function have been examined in association studies. Early studies on host genetic polymorphism and infectious disease focused primarily on human leukocyte antigen (HLA) genes, and included early examples of family-based association studies.

Selective pressure by infectious diseases is likely to account for the much higher levels of polymorphism seen in human HLA regions (BOX 1) when compared with other regions of the genome. There are examples of protective HLA types that seem to have been driven to high allele frequencies by infectious diseases[11], and polymorphisms are seen in HLA molecules predominantly at sites that are crucial for peptide EPITOPE binding[39] (FIG. 3).

High levels of polymorphism might be maintained through a type of frequency-dependent selection (in which pathogens have evolved so that their constituent peptides cannot be bound by the more prevalent HLA molecules). Alternatively, polymorphism might be maintained through balancing selection between diseases, in which alleles that are disadvantageous for one disease are advantageous for another, or through overdominant selection (heterozygote advantage)[40], in which heterozygotes are most resistant to disease. Some evidence now exists to support an advantage for HLA heterozygosity, which has been associated with protection against several viral diseases, including delayed progression of HIV infection[41] and resistance to hepatitis B virus (HBV) carriage[42].

The results of association studies on infectious disease often produce conflicting results. The potential reasons for this are common to association studies of other complex diseases[43]. First, many older studies used sample sizes that were too small to detect association reliably, and many putative associations lost significance after the appropriate adjustment was made for the number of comparisons actually undertaken. Second, studies could be confounded by POPULATION STRATIFICATION, which gives rise to false-positive associations. Confirming ethnicity in study groups is therefore very important. Third, an allele that shows association might be in LINKAGE DISEQUILIBRIUM (LD) with another functional polymorphism. Futhermore, patterns of LD vary between populations, so two seemingly different associations might reflect an underlying common allelic association with a flanking locus. In addition, studies must contend with variation in pathogens between geographically separated areas, which leads to association with, and selection of, different HLA variants[44]. This is likely to be of greater importance for diseases such as malaria, in which pathogen variability is high, and of less importance in diseases such as tuberculosis, in which pathogen diversity is low.

Despite the limitations of some of the early studies, there is now a convincing list of diseases that show association with specific HLA alleles, and the list is likely to continue to expand (TABLE 3). For example, common West African HLA alleles show association with protection against severe malaria, providing evidence of the selective pressure from this infectious disease on HLA regions[44]. In particular, the protective effect of common class I HLA alleles has encouraged efforts to develop liver-stage anti-malarial vaccines (FIG. 2) because hepatocytes, but not erythrocytes, display class I molecules.

HLA associations with HIV progression have also been investigated in some detail. Some specific class I HLA types, such as B27 and B57, have been associated with a better prognosis[45] and others, including allelic variants of B35, with poorer prognoses[41,46,47]. Recently, it has been shown that a single amino-acid change can make the difference between a neutral and a deleterious allele[48]. Convincing HLA associations with HIV progression correlate well with our understanding of the interaction between the pathogen and the host[49].

HLA associations have been found consistently between mycobacterial diseases and class II alleles, particularly HLA-DR subtypes. Associations with tuberculosis and both tuberculoid and lepromatous leprosy (see TABLE 1) have been established clearly in India[50] and elsewhere[51]. HLA class II alleles have also shown confirmed association with hepatitis clearance (reviewed in REF. 52). The HLA-DRB1*1302 allele is associated with spontaneous clearance of HBV infection in The Gambia, Europe and Korea. Several studies have shown the HLA-DRB1*1101–DQB1*0301 haplotype to be more frequent in those who clear hepatitis C virus infection[53] (one of the most consistent associations of an HLA type with an infectious disease). Reproducible associations with infection and HLA molecules have led to considerable interest in identifying the pathogen epitopes that they present to the immune system. Several such epitopes are now part of subunit-based vaccines that are undergoing evaluation in clinical trials.

EPITOPE
The portion of an antigen that interfaces with the antigen-binding site of an antibody or T-cell receptor.

POPULATION STRATIFICATION
A population that contains several sub-populations that differ in their genetic characteristics.

LINKAGE DISEQUILIBRIUM
When the frequency of a particular haplotype for two loci is significantly greater than that expected from the product of the observed allelic frequencies at each locus.

In malaria, tuberculosis and probably other infections, antigen-specific immune responses show varying degrees of heritability and of HLA restriction. There is evidence that HLA genes might only account for a small proportion of the heritability for some antigen-specific responses, with non-HLA genes having a greater effect[54,55]. This is in contrast to some autoimmune diseases, such as type I diabetes, in which the HLA region accounts for a larger proportion of genetically determined disease susceptibility[56].

Non-HLA candidate-gene studies

The assessment of candidate genes in infectious-disease association studies has provided important evidence for the clinical relevance of *in vitro* observations in particular diseases, as well as important insights into disease pathogenesis.

Malaria. After the success in identifying mutations that protect against malaria, many of the polygenes that are important for susceptibility to malaria are now being identified. Comparing large numbers of patients affected by severe malaria (cerebral malaria, hypoglycaemia or severe malarial anaemia) with appropriate controls has proved the most successful strategy for identifying protective mutations. Variation has been studied in several genes that affect correlates of disease severity, such as adhesion to the blood vessel wall and ROSETTING. Tumour-necrosis factor-α (TNF-α) — a CYTOKINE that is secreted from white blood cells with widespread effects on immune activation — has been the subject of considerable investigation. Clinical studies have shown correlations between high TNF-α levels in the serum and forms of severe malaria. Genetic studies are beginning to uncover mechanisms that are responsible for TNF-α-mediated effects, and two independent TNF-α promoter polymorphisms have been associated with severe disease[57]. One of these polymorphisms, TNF-α-308, might cause increased expression of TNF-α[58]. Another one influences transcription-factor binding[59] and is associated with an increased risk of cerebral malaria.

CD36 has a role in the endothelial adhesion of infected erythrocytes and their clearance by the immune system. The observation that CD36 deficiency is common in Africa prompted investigations into the role of this molecule in malaria. CD36 deficiency has recently been associated with both susceptibility and resistance to forms of severe malaria[60,61], which indicates that spatial and even temporal variation in parasite polymorphism might contribute to heterogeneity in disease associations.

Geographical variation in the genetic associations of the host with disease is now well documented in malarial studies, perhaps reflecting the extensive polymorphism of *P. falciparum*, as well as different dominant parasite epitopes found in different geographical locations. For example, ICAM[Kilifi], which is a variant of ICAM1 (intercellular adhesion molecule 1 (CD54)) — a general adhesion molecule — is found more commonly in Kenyan children with severe malaria[62], but is not associated with disease in West Africans[63].

Mycobacterial infection. Although both are caused by mycobacteria, leprosy and tuberculosis differ greatly in their clinical phenotypes. However, some similarities are being detected in the role of certain candidate genes, in particular *NRAMP1* (recently renamed *SLC11A1* — solute carrier family 11, member 1) and the vitamin D receptor gene (*VDR*).

NRAMP1 was identified by several groups working on a mouse locus that confers susceptibility to intracellular infections, such as *Leishmania*, *Salmonella* and the BCG strain of *Mycobacterium bovis*[64]. NRAMP1, like the related NRAMP2 (SLC11A2), is probably a divalent cation transporter and is found in the membrane of the PHAGOLYSOSOMES[65]. In mouse models, *Nramp1* is important in resistance to several intracellular infections. Although, in humans, the effect of *NRAMP1* variants seems more modest, association has been found between tuberculosis and *NRAMP1* in populations as diverse as West Africans[66,67], Japanese and Koreans[68]. Linkage between tuberculosis and the *NRAMP1* locus has been shown in a large Canadian pedigree[69], but linkage was not seen in Brazilian, West African or South African populations[70,71]. Although weak evidence of linkage between the *NRAMP1* locus and leprosy has been reported, it was not confirmed[72]. However, a subsequent Malian study again found an association between *NRAMP1* and leprosy[73].

Apart from its role in calcium and bone metabolism, the active metabolite of vitamin D (1,25-OH D3) is important in modulating immune responses. Among its many actions, it suppresses lymphocyte proliferation, immunoglobulin production and cytokine synthesis; it also stimulates cell-mediated immunity, alters dendritic cell maturation and activates macrophages to inhibit *M. tuberculosis* growth *in vitro*[74]. These effects are brought about through the VDR, which is present in macrophages, and both B and T lymphocytes. Evidence that supports a role for vitamin D in tuberculosis includes the clinical observation that deficiency is probably associated with increased risk of tuberculosis[75], and seasonal variation in vitamin D levels has been suggested as the explanation of the seasonal incidence pattern of tuberculosis[76]. A polymorphism that alters a *Taq*I restriction site in VDR has been associated

Table 3 | **Associations between HLA alleles and infectious disease**

	Disease	Effect	Reference
Class I			
B8	Pulmonary tuberculosis	Susceptibility	50
B35	AIDS	Susceptibility	41
B53	Severe malaria	Resistance	11
B57	AIDS	Resistance	45
Class II			
DRB1*1302	Hepatitis B	Resistance	120
DRB1*1352	Malarial anaemia	Resistance	11
DRB1*1101	Hepatitis C	Resistance	121
DRB1*04	Typhoid fever	Resistance	122
DR2	Pulmonary tuberculosis	Susceptibility	50
DR2	Leprosy	Susceptibility	51
DR7	Hepatitis B	Susceptibility	123

HLA, human leukocyte antigen.

with the two polar types of leprosy found in India[77] and with resistance to tuberculosis[75,78].

Bacterial infection. Despite few estimates of heritability of acute bacterial infections there has been some success in identifying susceptibility loci for invasive bacterial infection using the candidate-gene approach. The best example is provided by mannose-binding LECTIN (MBL2), which is a C-type (calcium-dependent) serum lectin with a role in INNATE IMMUNITY. It has a role in the OPSONIZATION of bacteria by binding to ordered arrays of surface mannose and amino-acetylglucosamine residues on pathogen surfaces. It also effects complement activation by interaction with MBL-associated serum proteases (MASP1 and MASP2, mannan-binding lectin serine proteases)[79]. Many healthy individuals have very low MBL levels as a result of one of three single amino-acid changes, usually referred to by the affected codon — 52, 54 and 57. In most populations, these variants reach polymorphic frequencies[80] and ~5% of individuals are homozygous for non-functional alleles. It was originally suggested that low levels of MBL were associated with recurrent infections in children[81], and one study of ethnically diverse children from London found that mutant alleles were more common in those admitted to hospital with infections than in non-infectious controls[82]. In Greenland, the same alleles were associated with an increased risk of respiratory infection before the age of 2 (REF. 83). In larger studies, association with specific infectious diseases has been harder to establish[84], although an excess of mutant homozygotes has been claimed in some cohorts with meningococcal disease[85]. Because of these conflicts, it is clear that further studies are required to clarify the role of MBL in infectious disease. Recent evidence indicates that MBL deficiency might be more important in individuals with some pre-existing immunodeficiency[86], fuelling interest in therapeutic MBL replacement as a way to prevent infection.

Chronic hepatitis. Only 10–20% of those infected with HBV progress to chronic infection. Early studies noted that carriage of hepatitis B surface antigen, which reflects persistent infection, sometimes seemed to be familial[87]. The first segregation analyses suggested the presence of an autosomal-recessive gene[88], although later studies supported a polygenic model. In contrast to hepatitis B, most people infected by hepatitis C go on to develop persistent infection.

IL10 polymorphisms seem to have a role in hepatitis, as well as in HIV progression[89]. IL10 is a key immunoregulatory cytokine that can inhibit T-cell proliferation and cytokine secretion[90]. Promoter polymorphisms have been found to be associated with the risk of hepatitis B persistence (L. Zhang *et al.*, unpublished observations) and altered response to IFN-α therapy in hepatitis C infection[91,92].

HIV/AIDS. A relatively new human infectious disease, such as HIV, might show stronger genetic associations because there has been insufficient time for markedly

deleterious alleles to be eliminated by natural selection. Studies of HIV have been based mainly on case–control methodologies and long-term cohort studies of disease progression. Long-term cohort analysis has been the preferred approach because of the difficulties in defining 'at risk' controls and has offered the possibility of following large numbers of patients.

The discovery that, along with CD4, chemokine receptors are crucial co-receptors for the entry of some HIV strains into macrophages and lymphocytes has led to studies into the effect of polymorphism in these genes on HIV infection and disease progression. The best-studied variant, a 32-bp deletion in the promoter of CCR5 (CC chemokine receptor 5), confers protection for homozygotes against HIV infection[93], and heterozygotes have a significant delay in the progression to AIDS and death[94]. These associations are supported by studies that describe the functional effects of the mutation[95].

The *CCR5Δ32* mutation is present in 8–21% (REF. 93) of Caucasian populations, but is not found in African or most Asian populations. It is found as part of a rare haplotype, suggesting recent selective pressure. However, given that HIV probably entered the population within the last 100 years, it is too recent to have been the selecting force[96]. Another infectious disease was the most probable selective agent, and some authors have speculated that *Yersinia pestis* — the causative agent of bubonic plague — might have been responsible.

More prevalent *CCR5* polymorphisms have also been studied in some detail. Homozygosity for a particular promoter polymorphism has been associated with transmission of HIV to African–American infants[97] and, in several studies, single nucleotide polymorphisms (SNPs), or haplotypes of these SNPs in the promoter of *CCR5*, have been associated with accelerated progression to AIDS[98]. Association has also been found and confirmed with the CCR2 Ile64 variant[99], and this might reflect its capacity to heterodimerize with CCR5. Another molecule of interest is CXC chemokine receptor 4 (CXCR4), which seems to be a co-receptor for viral types seen later in infection. Individuals who are homozygous for the 3′ untranslated region-A801 polymorphism in SDF1 (stromal cell-derived factor 1, a CXCR4 ligand that can downregulate CXCR4 expression) might have altered rates of progression to AIDS[100,101], but reports on this association are inconsistent.

Linkage analysis and immunogenomics
Most contributory genes identified so far account for a small proportion of genetic variability in susceptibility to infectious disease. As a result of developments in the genetic analysis of complex traits, genome-wide analysis of many hundreds of highly polymorphic markers can now be carried out to identify genomic regions that show linkage to disease. Such approaches, coupled with rapidly expanding databases of genome sequence and mapped genes, allow the identification of important genes on the basis of their genomic location, without necessarily any knowledge of their biological effects or their mode of inheritance. Such studies typically require the recruitment of numerous families with more than one affected child.

LECTIN
A sugar-binding receptor of the innate immune system.

INNATE IMMUNITY
That part of the immune response that is not adaptive (that is, does not change with repeated exposure to a pathogen).

OPSONIZATION
The process whereby molecules that are deposited on the surface of pathogens (opsonins) aid uptake of these pathogens into the specialized cells of the immune system.

There has been considerable investment in linkage-based studies for non-communicable diseases, and this research is beginning to yield interesting results, such as the recent identification of *NOD2* as a major susceptibility gene for Crohn disease[26,27]. The suitability of an infectious disease to such an approach depends on several factors, both theoretical and practical (such as the existence of reliable medical records). The most prevalent worldwide infectious diseases usually coexist with the poorest standards of living and the worst public health infrastructure. Even if such infrastructure exists, ascertainment and recruitment of numerous families can be difficult. Nonetheless, although the application of such large-scale approaches to infectious disease is relatively new and unattractive to commercial funding, marked progress has been made. Genome-wide analysis has identified regions that might contain major susceptibility genes[5,6,102]. The discovery of such loci is encouraging given the highly polygenic nature of infectious disease genetics and the relatively low sibling risks measured in comparison with some other complex diseases.

Perhaps the best example of a disease that is suitable for such a study is leprosy. Leprosy has a measured heritable component[103], is still common in many countries and is sufficiently INDOLENT to allow the collection of numerous affected families. The phenotype is also generally clear and visible, which helps to reduce bias in collecting families and, perhaps most importantly, the attributable mortality is relatively low, which indicates that if a major susceptibility gene does exist it might be able to persist in the population. Significant evidence for linkage has recently been identified between the chromosomal region 10p13 (REF. 5) and leprosy in India, which provides evidence that linkage studies can map major non-HLA loci for a human infectious disease. This locus might account for ~40% of the overall genetic component in the south Indian population studied. This result is encouraging for those involved in such studies as it allays initial concerns that the polygenic nature of infectious disease might make such mapping impossible.

In a genome-wide linkage study of tuberculosis, two genomic regions — one near the centromere of chromosome 15, the other at Xq27 — showed suggestive evidence of linkage with tuberculosis in African families[71]. Further support for the presence of susceptibility genes in these regions was provided by a new form of homozygosity mapping termed COMMON ANCESTRY MAPPING[71]. Interestingly, there is a well-documented excess of males with tuberculosis in Africa and an X-chromosomal major gene effect could contribute to this.

Chronic hepatitis infection is well suited to affected sibling-pair (sib-pair) linkage analysis, and searches for new genes that exert a major effect on this infection have had some success. Early results indicate that loci on chromosomes 6q and 21q show evidence of linkage and association with chronic HBV infection[104] (L. Zhang, unpublished data).

As well as looking at disease phenotype, similar approaches have been taken to examining infectious burden. Using affected sib-pair analysis, linkage has been shown between malarial parasite density and the chromosomal region 5q31–33 (REF. 105). A complete genome screen of siblings that are susceptible to high worm burdens of another parasite, *Schistosoma mansoni*, showed linkage to the same region (5q31–33)[6,102]. 5q31–33 contains genes for several immune-related molecules, notably cytokines, IL4, IL9 and IL13. This region has also shown linkage with phenotypes such as IgE level[106], ATOPY[107], asthma[108] and allergic dermatitis[109]. Linkage to the same region does not necessarily imply that the same alleles, or even the same genes, are contributing to these different diseases, but continuing studies should address this evolutionarily intriguing possibility.

Concluding remarks

Developments in modern genetics and genomics have contributed to our understanding of the pathogenic processes that underlie major infectious diseases by allowing a more systematic study of the genetic influences. The number of candidate susceptibility genes is expanding rapidly. A good example of this is the family of toll-like receptors. Initial studies of a mouse model with a reduced inflammatory response to ENDOTOXIN have identified causal mutations in Tlr4 (toll-like receptor 4). Exploration of the human genome has identified a family of similar molecules, and the importance of several of these in immune responses to specific infectious particles has recently been established[110]. Such molecules have become important candidates in the study of human infectious diseases.

Genome-wide linkage analysis is also beginning to provide insights into complex disease. Advances in SNP typing, microarray technology and bioinformatics will be as helpful in the study of infectious diseases as they are in non-infectious diseases. These developing technologies will allow genome-wide studies for disease genes to move beyond looking for linkage to association; efforts are being made to collect large cohorts so that small genetic effects can be detected.

What will be the practical benefits of this information? The finding of protective associations with particular HLA types has encouraged the development of new epitope-based vaccines and stimulated new areas of immunological research. The role of CCR5 variation in HIV infection has stimulated research into this molecule as a drug target. There is considerable interest in using genotyping to prioritize the use of expensive BIOLOGICS, such as IFN-α therapy in chronic hepatitis C virus infection, as it becomes possible to predict who might respond favourably to treatment. As more and more molecules that are involved in innate, as well as acquired, immunity are characterized, the possibility of being able to predict the risk profile of an individual for specific infectious diseases gets closer. Until recently, pathogen diversity has been seen as an important obstacle to studies in infectious susceptibility. As molecular typing allows the identification of related strains of bacteria, and virulence molecules are characterized in greater detail, it should be possible to tease out strain-specific host–pathogen interactions at the molecular level[111]. This will add to our understanding of the complexities of this interplay in the pathogenesis of major infectious diseases.

INDOLENT
In medical terms, slow to develop or heal.

COMMON ANCESTRY MAPPING
Method for finding genes for susceptibility to disease by identifying discrepancies in the time to most recent common ancestor in homologous regions from susceptible and non-susceptible individuals.

ATOPY
Clinical manifestations of allergic immune responses; for example, eczema or asthma.

ENDOTOXIN
Lipopolysaccharide that is found in the membrane of Gram-negative bacteria and that activates B cells and macrophages.

BIOLOGICS
Agents of biological origin that are used to diagnose or treat disease.

1. James, S. P., Nicol, W. D. & Shute, P. G. A study of malignant tertian malaria. *Proc. R. Soc. Med.* **25**, 1153 (1932).

2. Modiano, D. *et al.* Different response to *Plasmodium falciparum* malaria in west African sympatric ethnic groups. *Proc. Natl Acad. Sci. USA* **93**, 13206–13211 (1996).
   Study of three West African ethnic groups with similar exposure to malaria. It provides evidence that one group, the Fulani, might have genetic factors that enhance their immunity to malaria.

3. Stead, W. W., Senner, J. W., Reddick, W. T. & Lofgren, J. P. Racial differences in susceptibility to infection by *Mycobacterium tuberculosis*. *N. Engl. J. Med.* **322**, 422–427 (1990).

4. Sorensen, T. I., Nielsen, G. G., Andersen, P. K. & Teasdale, T. W. Genetic and environmental influences on premature death in adult adoptees. *N. Engl. J. Med.* **318**, 727–732 (1988).

5. Siddiqui, M. R. *et al.* A major susceptibility locus for leprosy in India maps to chromosome 10p13. *Nature Genet.* **27**, 439–441 (2001).
   The first demonstration that a major disease locus can be mapped using genome-wide linkage analysis in a complex infectious disease.

6. Marquet, S. *et al.* Genetic localization of a locus controlling the intensity of infection by *Schistosoma mansoni* on chromosome 5q31–q33. *Nature Genet.* **14**, 181–184 (1996).

7. McGuffin, P. & Huckle, P. Simulation of Mendelism revisited: the recessive gene for attending medical school. *Am. J. Hum. Genet.* **46**, 994–999 (1990).

8. Haldane, J. B. S. Disease and evolution. *Ric. Sci.* (Suppl. A), 68–76 (1949).

9. Lederberg, J. J. B. S. Haldane (1949) on infectious disease and evolution. *Genetics* **153**, 1–3 (1999).

10. Allison, A. C. Protection afforded by sickle-cell trait against subtertian malarial infection. *Br. Med. J.* 290–294 (1954).

11. Hill, A. *et al.* Common West African HLA antigens are associated with protection from severe malaria. *Nature* **352**, 595–600 (1991).

12. Pasvol, G., Weatherall, D. J. & Wilson, R. J. Cellular mechanism for the protective effect of haemoglobin S against *P. falciparum* malaria. *Nature* **274**, 701–703 (1978).

13. Weatherall, D. J. Phenotype–genotype relationships in monogenic disease: lessons from the thalassaemias. *Nature Rev. Genet.* **2**, 245–255 (2001).

14. Flint, J. *et al.* High frequencies of α-thalassaemia are the result of natural selection by malaria. *Nature* **321**, 744–750 (1986).

15. Willcox, M. *et al.* A case–control study in northern Liberia of *Plasmodium falciparum* malaria in haemoglobin S and β-thalassaemia traits. *Ann. Trop. Med. Parasitol.* **77**, 239–246 (1983).

16. Williams, T. N. *et al.* High incidence of malaria in α-thalassaemic children. *Nature* **383**, 522–525 (1996).
   Illustrates the new insights that are being gained into malarial immunity through the study of monogenic traits.

17. Ruwende, C. *et al.* Natural selection of hemi- and heterozygotes for G6PD deficiency in Africa by resistance to severe malaria. *Nature* **376**, 246–249 (1995).

18. Friedman, M. J. Oxidant damage mediates variant red cell resistance to malaria. *Nature* **280**, 245–247 (1979).

19. Levine, M. M. *et al.* Genetic susceptibility to cholera. *Ann. Hum. Biol.* **6**, 369–374 (1979).

20. Tournamille, C., Colin, Y., Cartron, J. P. & Le Van Kim, C. Disruption of a GATA motif in the *Duffy* gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nature Genet.* **10**, 224–228 (1995).

21. Miller, L. H., Mason, S. J., Clyde, D. F. & McGinniss, M. H. The resistance factor to *Plasmodium vivax* in blacks. The Duffy-blood-group genotype, FyFy. *N. Engl. J. Med.* **295**, 302–304 (1976).

22. Bertranpetit, J. & Calafell, F. Genetic and geographical variability in cystic fibrosis: evolutionary considerations. *Ciba Found. Symp.* **197**, 97–114 (1996).
   Extensive discussion of the evidence that the gene responsible for cystic fibrosis has been subject to natural selection.

23. Thompson, E. A. & Neel, J. V. Allelic disequilibrium and allele frequency distribution as a function of social and demographic history. *Am. J. Hum. Genet.* **60**, 197–204 (1997).

24. Pier, G. B. *et al.* Role of mutant CFTR in hypersusceptibility of cystic fibrosis patients to lung infections. *Science* **271**, 64–67 (1996).

25. Pier, G. B. *et al. Salmonella typhi* uses CFTR to enter intestinal epithelial cells. *Nature* **393**, 79–82 (1998).

26. Hugot, J. P. *et al.* Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* **411**, 599–603 (2001).

27. Ogura, Y. *et al.* A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* **411**, 603–606 (2001).
   References 26 and 27 describe two different approaches that show that major susceptibility genes can be mapped through linkage analysis of complex disease.

28. Fischer, A. Primary immunodeficiency disease: an experimental model for molecular medicine. *Lancet* **357**, 1863–1869 (2001).

29. Korthauer, U. *et al.* Defective expression of T-cell CD40 ligand causes X-linked immunodeficiency with hyper-IgM. *Nature* **361**, 539–541 (1993).

30. DiSanto, J. P., Bonnefoy, J. Y., Gauchat, J. F., Fischer, A. & de Saint Basile, G. CD40 ligand mutations in X-linked immunodeficiency with hyper-IgM. *Nature* **361**, 541–543 (1993).

31. Newport, M. J. *et al.* A mutation in the interferon-γ-receptor gene and susceptibility to mycobacterial infection. *N. Engl. J. Med.* **335**, 1941–1949 (1996).

32. Jouanguy, E. *et al.* A human IFNGR1 small deletion hotspot associated with dominant susceptibility to mycobacterial infection. *Nature Genet.* **21**, 370–378 (1999).

33. Altare, F. *et al.* Inherited interleukin 12 deficiency in a child with bacille Calmette–Guerin and *Salmonella enteritidis* disseminated infection. *J. Clin. Invest.* **102**, 2035–2040 (1998).

34. De Jong, R. *et al.* Severe mycobacterial and *Salmonella* infections in interleukin-12 receptor-deficient patients. *Science* **280**, 1435–1438 (1998).

35. Campbell, S. J. The identification of genetic susceptibility factors for tuberculosis. Ph.D. thesis, University of Oxford, UK (2001).

36. Ting, L. M., Kim, A. C., Cattamanchi, A. & Ernst, J. D. *Mycobacterium tuberculosis* inhibits IFN-γ transcriptional responses without inhibiting activation of STAT1. *J. Immunol.* **163**, 3898–3906 (1999).

37. Dupuis, S. *et al.* Impairment of mycobacterial but not viral immunity by a germline human *STAT1* mutation. *Science* **293**, 300–303 (2001).

38. Murphy, P. M. Molecular mimicry and the generation of host defense protein diversity. *Cell* **72**, 823–826 (1993).

39. Hughes, A. L. & Nei, M. Nucleotide substitution at major histocompatability complex class II loci: evidence for overdominant selection. *Proc. Natl Acad. Sci. USA* **86**, 958–962 (1989).

40. Doherty, P. C. & Zinkernagel, R. M. The biological role of the major histocompatibility antigens. *Lancet* **1**, 1406–1409 (1975).

41. Carrington, M. *et al.* HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science* **283**, 1748–1752 (1999).

42. Thursz, M. R., Thomas, H. C., Greenwood, B. M. & Hill, A. V. Heterozygote advantage for HLA class-II type in hepatitis B virus infection. *Nature Genet.* **17**, 11–12 (1997).

43. Cardon, L. R. & Bell, J. I. Association study designs for complex diseases. *Nature Rev. Genet.* **2**, 91–99 (2001).
   Extensive review of the design and interpretation of association studies in complex disease.

44. Hill, A. V. HIV and HLA: confusion or complexity? *Nature Med.* **2**, 395–396 (1996).

45. Kaslow, R. A. *et al.* Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection. *Nature Med.* **2**, 405–411 (1996).

46. Sahmoud, T. *et al.* Progression to AIDS in French haemophiliacs: association with HLA-B35. *Aids* **7**, 497–500 (1993).

47. Scorza Smeraldi, R. *et al.* HLA-associated susceptibility to acquired immunodeficiency syndrome in Italian patients with human-immunodeficiency-virus infection. *Lancet* **2**, 1187–1189 (1986).

48. Gao, X. *et al.* Effect of a single amino acid change in MHC class I molecules on the rate of progression to AIDS. *N. Engl. J. Med.* **344**, 1668–1675 (2001).

49. Kelleher, A. D. *et al.* Clustered mutations in HIV-1 gag are consistently required for escape from HLA-B27-restricted cytotoxic T lymphocyte responses. *J. Exp. Med.* **193**, 375–386 (2001).

50. Brahmajothi, V. *et al.* Association of pulmonary tuberculosis and HLA in south India. *Tubercle* **72**, 123–132 (1991).

51. Visentainer, J. E., Tsuneto, L. T., Serra, M. F., Peixoto, P. R. & Petzl-Erler, M. L. Association of leprosy with HLA-DR2 in a Southern Brazilian population. *Braz. J. Med. Biol. Res.* **30**, 51–59 (1997).

52. Thursz, M. MHC and the viral hepatitides. *Quart. J. Med.* **94**, 287–291 (2001).

53. Alric, L. *et al.* Genes of the major histocompatibility complex class II influence the outcome of hepatitis C virus infection. *Gastroenterology* **113**, 1675–1681 (1997).

54. Jepson, A. *et al.* Quantification of the relative contribution of major histocompatibility complex (MHC) and non-MHC genes to human immune responses to foreign antigens. *Infect. Immun.* **65**, 872–876 (1997).

55. Jepson, A. *et al.* Genetic regulation of acquired immune responses to antigens of *Mycobacterium tuberculosis*: a study of twins in West Africa. *Infect. Immun.* **69**, 3989–3994 (2001).

56. Davies, J. L. *et al.* A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* **371**, 130–136 (1994).

57. McGuire, W. *et al.* Severe malarial anemia and cerebral malaria are associated with different tumor necrosis factor promoter alleles. *J. Infect. Dis.* **179**, 287–290 (1999).

58. Wilson, A. G., Symons, J. A., McDowell, T. L., McDevitt, H. O. & Duff, G. W. Effects of a polymorphism in the human tumor necrosis factor α promoter on transcriptional activation. *Proc. Natl Acad. Sci. USA* **94**, 3195–3199 (1997).

59. Knight, J. C. *et al.* A polymorphism that affects OCT-1 binding to the TNF promoter region is associated with severe malaria. *Nature Genet.* **22**, 145–150 (1999).
   Important evidence of the functional properties of the tumour necrosis factor promoter polymorphisms that are associated with severe malaria.

60. Aitman, T. J. *et al.* Malaria susceptibility and CD36 mutation. *Nature* **405**, 1015–1016 (2000).

61. Pain, A. *et al.* A non-sense mutation in *Cd36* gene is associated with protection from severe malaria. *Lancet* **357**, 1502–1503 (2001).
   References 60 and 61 show the heterogeneity of data in genetic studies and how such data can inform research on particular candidate genes.

62. Fernandez-Reyes, D. *et al.* A high frequency African coding polymorphism in the N-terminal domain of *ICAM-1* predisposing to cerebral malaria in Kenya. *Hum. Mol. Genet.* **6**, 1357–1360 (1997).

63. Bellamy, R., Kwiatkowski, D. & Hill, A. V. Absence of an association between intercellular adhesion molecule 1, complement receptor 1 and interleukin 1 receptor antagonist gene polymorphisms and severe malaria in a West African population. *Trans. R. Soc. Trop. Med. Hyg.* **92**, 312–316 (1998).

64. Vidal, S. M., Malo, D., Vogan, K., Skamene, E. & Gros, P. Natural resistance to infection with intracellular parasites: isolation of a candidate for BCG. *Cell* **73**, 469–485 (1993).

65. Gruenheid, S. & Gros, P. Genetic susceptibility to intracellular infections: *Nramp1*, macrophage function and divalent cations transport. *Curr. Opin. Microbiol.* **3**, 43–48 (2000).

66. Bellamy, R. *et al.* Variations in the *NRAMP1* gene and susceptibility to tuberculosis in West Africans. *N. Engl. J. Med.* **338**, 640–644 (1998).

67. Cervino, A. C., Lakiss, S., Sow, O. & Hill, A. V. Allelic association between the *NRAMP1* gene and susceptibility to tuberculosis in Guinea-Conakry. *Ann. Hum. Genet.* **64**, 507–512 (2000).

68. Gao, P. S. *et al.* Genetic variants of *NRAMP1* and active tuberculosis in Japanese populations. International Tuberculosis Genetics Team. *Clin. Genet.* **58**, 74–76 (2000).

69. Greenwood, C. M. *et al.* Linkage of tuberculosis to chromosome 2q35 loci, including *NRAMP1*, in a large aboriginal Canadian family. *Am. J. Hum. Genet.* **67**, 405–416 (2000).

70. Shaw, M. A. *et al.* Evidence that genetic susceptibility to *Mycobacterium tuberculosis* in a Brazilian population is under oligogenic control: linkage study of the candidate genes *NRAMP1* and *TNFA*. *Tuber. Lung Dis.* **78**, 35–45 (1997).

71. Bellamy, R. *et al.* Genetic susceptibility to tuberculosis in Africans: a genome-wide scan. *Proc. Natl Acad. Sci. USA* **97**, 8005–8009 (2000).

72. Abel, L. *et al.* Susceptibility to leprosy is linked to the human *NRAMP1* gene. *J. Infect. Dis.* **177**, 133–145 (1998).

73. Meisner, S. J. *et al.* Association of *NRAMP1* polymorphism with leprosy type but not susceptibility to leprosy *per se* in West Africans. *Am. J. Trop. Med. Hyg.* (in the press).

74. Rook, G. A. *et al.* Vitamin D3, γ interferon, and control of proliferation of *Mycobacterium tuberculosis* by human monocytes. *Immunology* **57**, 159–163 (1986).

75. Wilkinson, R. J. *et al.* Influence of vitamin D deficiency and vitamin D receptor polymorphisms on tuberculosis among Gujarati Asians in west London: a case–control study. *Lancet* **355**, 618–621 (2000).

76. Douglas, A. S., Strachan, D. P. & Maxwell, J. D. Seasonality of tuberculosis: the reverse of other respiratory diseases in the UK. *Thorax* **51**, 944–946 (1996).

77. Roy, S. *et al.* Association of vitamin D receptor genotype with leprosy type. *J. Infect. Dis.* **179**, 187–191 (1999).

78. Bellamy, R. *et al*. Tuberculosis and chronic hepatitis B virus infection in Africans and variation in the vitamin D receptor gene. *J. Infect. Dis.* **179**, 721–724 (1999).
79. Thiel, S. *et al*. Interaction of C1q and mannan-binding lectin (MBL) with C1r, C1s, MBL-associated serine proteases 1 and 2, and the MBL-associated protein MAp19. *J. Immunol.* **165**, 878–887 (2000).
80. Madsen, H. O. *et al*. A new frequent allele is the missing link in the structural polymorphism of the human mannan-binding protein. *Immunogenetics* **40**, 37–44 (1994).
81. Summerfield, J. A. *et al*. Mannose binding protein gene mutations associated with unusual and severe infections in adults. *Lancet* **345**, 886–889 (1995).
82. Summerfield, J. A., Sumiya, M., Levin, M. & Turner, M. W. Association of mutations in mannose binding protein gene with childhood infection in consecutive hospital series. *Br. Med. J.* **314**, 1229–1232 (1997).
83. Koch, A. *et al*. Acute respiratory tract infections and mannose-binding lectin insufficiency during early childhood. *J. Am. Med. Assoc.* **285**, 1316–1321 (2001).
84. Bellamy, R. *et al*. Mannose binding protein deficiency is not associated with malaria, hepatitis B carriage nor tuberculosis in Africans. *Quart. J. Med.* **91**, 13–18 (1998).
85. Hibberd, M. L., Sumiya, M., Summerfield, J. A., Booy, R. & Levin, M. Association of variants of the gene for mannose-binding lectin with susceptibility to meningococcal disease. Meningococcal Research Group. *Lancet* **353**, 1049–1053 (1999).
86. Neth, O., Hann, I., Turner, M. W. & Klein, N. J. Deficiency of mannose-binding lectin and burden of infection in children with malignancy: a prospective study. *Lancet* **358**, 614–618 (2001).
87. Lin, H. J. *et al*. Evidence for intrafamilial transmission of hepatitis B virus from sequence analysis of mutant HBV DNAs in two Chinese families. *Lancet* **336**, 208–212 (1990).
88. Blumberg, B. S., Friedlander, J. S., Woodside, A., Sutnick, A. I. & London, W. T. Hepatitis and Australia antigen: autosomal recessive inheritance of susceptibility to infection in humans. *Proc. Natl Acad. Sci. USA* **62**, 1108–1115 (1969).
89. Shin, H. D. *et al*. Genetic restriction of HIV-1 pathogenesis to AIDS by promoter alleles of IL10. *Proc. Natl Acad. Sci. USA* **97**, 14467–14472 (2000).
90. Akdis, C. A. & Blaser, K. Mechanisms for interleukin-10 mediated immune suppression. *Immunology* **103**, 131–136 (2001).
91. Edwards-Smith, C. J. *et al*. Interleukin-10 promoter polymorphism predicts initial response of chronic hepatitis C to interferon α. *Hepatology* **30**, 526–530 (1999).
92. Yee, L. J. *et al*. Interleukin 10 polymorphisms as predictors of sustained response in antiviral therapy for chronic hepatitis C infection. *Hepatology* **33**, 708–712 (2001).
93. Dean, M. *et al*. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the *CKR5* structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. *Science* **273**, 1856–1862 (1996).
94. Huang, Y. *et al*. The role of a mutant *CCR5* allele in HIV-1 transmission and disease progression. *Nature Med.* **2**, 1240–1243 (1996).
95. Benkirane, M., Jin, D. Y., Chun, R. F., Koup, R. A. & Jeang, K. T. Mechanism of transdominant inhibition of CCR5-

mediated HIV-1 infection by ccr5δ32. *J. Biol. Chem.* **272**, 30603–30606 (1997).
96. Zimmerman, P. A. *et al*. Inherited resistance to HIV-1 conferred by an inactivating mutation in CC chemokine receptor 5: studies in populations with contrasting clinical phenotypes, defined racial background, and quantified risk. *Mol. Med.* **3**, 23–36 (1997).
97. Kostrikis, L. G. *et al*. A polymorphism in the regulatory region of the CC-chemokine receptor 5 gene influences perinatal transmission of human immunodeficiency virus type 1 to African–American infants. *J. Virol.* **73**, 10264–10271 (1999).
98. Martin, M. P. *et al*. Genetic acceleration of AIDS progression by a promoter variant of *CCR5*. *Science* **282**, 1907–1911 (1998).
99. Smith, M. W. *et al*. Contrasting genetic influence of *CCR2* and *CCR5* variants on HIV-1 infection and disease progression. Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC), ALIVE Study. *Science* **277**, 959–965 (1997).
100. Winkler, C. *et al*. Genetic restriction of AIDS pathogenesis by an *SDF-1* chemokine gene variant. ALIVE Study, Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort (SFCC). *Science* **279**, 389–393 (1998).
101. Mummidi, S. *et al*. Genealogy of the *CCR5* locus and chemokine system gene variants associated with altered rates of HIV-1 disease progression. *Nature Med.* **4**, 786–793 (1998).
102. Marquet, S., Abel, L., Hillaire, D. & Dessein, A. Full results of the genome-wide scan which localises a locus controlling the intensity of infection by *Schistosoma mansoni* on chromosome 5q31–q33. *Eur. J. Hum. Genet.* **7**, 88–97 (1999).
   A report of the first genome-wide screen that used affected sib-pair analysis to study infectious burden.
103. Chakravarti, M. R. & Vogel, F. *A Twin Study on Leprosy* Vol. 1 (Stutthart Thieme, 1973).
104. Frodsham, A. The genetics of susceptibility to chronic hepatitis B infection. Ph.D. thesis, University of Oxford, UK (2000).
105. Rihet, P. *et al*. Malaria in humans: *Plasmodium falciparum* blood infection levels are linked to chromosome 5q31–q33. *Am. J. Hum. Genet.* **63**, 498–505 (1998).
106. Marsh, D. G. *et al*. Linkage analysis of *IL4* and other chromosome 5q31.1 markers and total serum immunoglobulin E concentrations. *Science* **264**, 1152–1156 (1994).
107. Walley, A. J., Wiltshire, S., Ellis, C. M. & Cookson, W. O. Linkage and allelic association of chromosome 5 cytokine cluster genetic markers with atopy and asthma associated traits. *Genomics* **72**, 15–20 (2001).
108. Holloway, J. W. *et al*. Linkage analysis of the 5q31–33 candidate region for asthma in 240 UK families. *Genes Immun.* **2**, 20–24 (2001).
109. Beyer, K. *et al*. Association and linkage of atopic dermatitis with chromosome 13q12–14 and 5q31–33 markers. *J. Invest. Dermatol.* **115**, 906–908 (2000).
110. Akira, S., Takeda, K. & Kaisho, T. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nature Immunol.* **2**, 675–680 (2001).

111. Gilbert, S. C. *et al*. Association of malaria parasite population structure, HLA, and immunological antagonism. *Science* **279**, 1173–1177 (1998).
   Demonstration of the complex interaction between variation in the human immune system and parasite variation in malaria.
112. Janeway, C. A. Jr & Travers, P. (eds) in *Immunobiology: the Immune System in Health and Disease*, 2nd Edn, Sect. 4 (1996).
113. Good, M. F. Towards a blood-stage vaccine for malaria: are we following all the leads? *Nature Rev. Immunol*. **1**, 117–125 (2001).
114. Smith, K. J. *et al*. Bound water structure and polymorphic amino acids act together to allow the binding of different peptides to MI+C class II HLA-B53. *Immunity* **4**, 215–228 (1996).
115. Diehl, K. & Von Verschuer, O. Der Erbeinfluss bei den Tuberkulose. *Beitr Klin Kunsch* **92**, 275 (1936).
116. Kallmann, F. J. & Reisner, D. Twin studies on the significance of genetic factors in tuberculosis. *Am. Rev. Respir. Dis.* **47**, 549 (1942).
117. Comstock, G. W. Tuberculosis in twins: a re-analysis of the Prophit survey. *Am. Rev. Respir. Dis.* **117**, 621 (1978).
118. Herndon, C. N. & Jennings, R. G. A twin family study on susceptibiltiy to poliomyelitis. *Am. J. Hum. Genet.* **3**, 17 (1951).
119. Lin, T. M. *et al*. Hepatitis B virus markers in Chinese twins. *Anticancer Res.* **9**, 737–741 (1989).
120. Thursz, M. R. *et al*. Association between an MHC class II allele and clearance of hepatitis B virus in the Gambia. *N. Engl. J. Med.* **332**, 1065–1069 (1995).
121. Thursz, M., Yallop, R., Goldin, R., Trepo, C. & Thomas, H. C. Influence of MHC class II genotype on outcome of infection with hepatitis C virus. The HENCORE group. Hepatitis C European Network for Cooperative Research. *Lancet* **354**, 2119–2124 (1999).
122. Dunstan, S. J. *et al*. Genes of the class II and class III major histocompatibility complex are associated with typhoid fever in Vietnam. *J. Infect. Dis.* **183**, 261–268 (2001).
123. Almarri, A. & Batchelor, J. R. HLA and hepatitis B infection. *Lancet* **344**, 1194–1195 (1994).

⊕ Online links

THE MAKING OF THE SOMITE: MOLECULAR EVENTS IN VERTEBRATE SEGMENTATION

Yumiko Saga and Hiroyuki Takeda

An incorrect version of Figure 4b was printed in this Review. The correct version is shown here, and this figure has been corrected in the full-text and PDF versions online. For the PDF see http://www.nature.com/cgi-taf/DynaPage.taf?file=/nrg/journal/v2/n11/full/nrg1101-835a_fs.html&filetype=pdf