Immunology

# Immunogenetics and genomics

*Adrian V S Hill*

**Immunogenetic analysis of disease susceptibility has been encouraged by the identification of strong HLA associations with several diseases of uncertain cause. Weaker HLA associations exist with a large number of infectious and non-infectious diseases and the mechanisms of these effects are beginning to be uncovered. Extensive analyses of non-HLA immunogenetic variants have also been undertaken and associations with a variety of genes identified. Genetic linkage analysis of multicase families has recently identified new major susceptibility loci for a few immunologically determined common diseases. However, the greatest potential for the future lies in genome-wide searches for susceptibility genes that individually might have quite modest effects but cumulatively have a large impact on individual risk. This new era of immunogenomics promises to provide key insights into disease pathogenesis and identify multiple molecular targets for intervention strategies.**

The explosion of genetic data emerging from genomics and high-throughput sequencing is generating unprecedented opportunities for defining the genetic basis of susceptibility to complex diseases. Although most of the important monogenic disorders have yielded to the power of modern molecular approaches, progress in unravelling the genetic basis of common multifactorial diseases has been slower. Simple evolutionary considerations suggest that the genetic basis of diseases involving the immune system might turn out to be the most complex of all, not just because of the many relevant genes and interpopulation differences, but because of the allelic diversity of these loci.

Immunogenetic analysis of human diseases developed alongside characterisation of the human leucocyte antigens, and some of the most striking immunogenetic associations still relate to this complex. The availability of hundreds of informative microsatellite markers spanning the human genome in the mid-1990s made possible searches for major susceptibility loci in family studies and started the current era of genome-wide studies. The genetic complexity of many diseases revealed by this approach has led to the view that a fuller understanding of polygenic disease will require hundreds of thousands of genetic markers. These would allow the entire genome to be examined by genetic association studies, rather than by linkage analyses that require multicase familes.[1] Although reservations remain about the usefulness of this new strategy, it is rapidly becoming feasible. An international consortium to identify more than a million single nucleotide polymorphisms (SNPs) is making good progress and several technologies for very rapid SNP typing are in development.

In this article, I will review some recent progress in the immunogenetics of infectious and non-infectious diseases.

## Measurement of genetic susceptibility

Efforts to identify genes for susceptibility to many autoimmune and infectious diseases have been advanced by studies of twins. Comparison of concordance rates in monozygotic and dizygotic twins provides an estimate of the size of the genetic component of susceptibility. For many infectious and autoimmune diseases, this component is substantial. A few investigators have also estimated the size of the genetic component of variable immune responses and estimated the overall effect of major histocompatibility complex (MHC) and non-MHC genes.[2] A frequently used measure of the size of the genetic component of disease susceptibility is the sibling risk, which is defined as the ratio of the disease risk in a sibling of an affected individual to the general population risk. For example, the sibling risk is 15 for type I diabetes and 20 for multiple sclerosis.[3] For infectious, and probably some non-infectious diseases, this value incorporates an environmental component caused by the greater environment sharing of siblings compared with the general population. However, the measure has been useful in providing an estimate of the importance of genetic factors, and can be compared to the locus-specific sibling risk values, derived from linkage studies, that provide estimates of how important individual loci are.

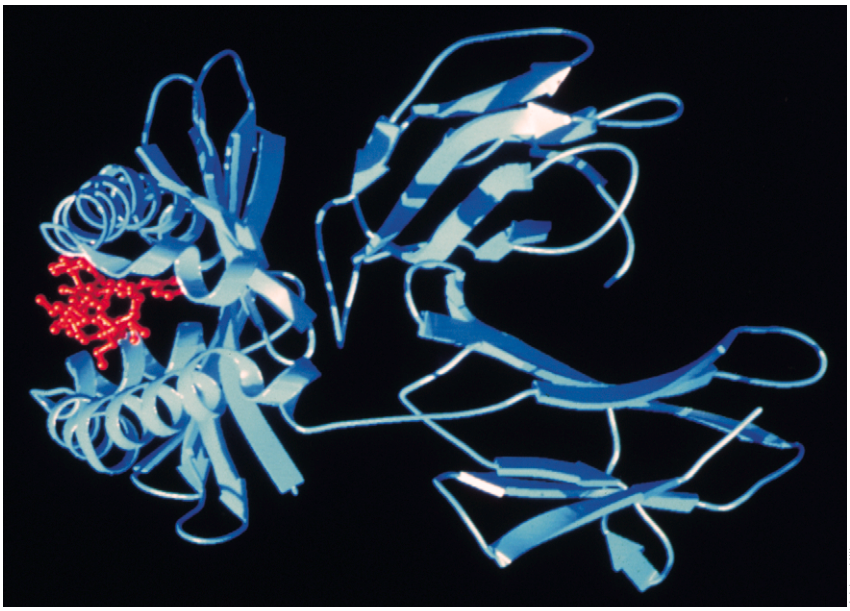
The results of studies of autoimmune and infectious diseases indicate that much of the genetic component is specified by many minor susceptibility genes, sometimes termed polygenes, rather than by a few major loci. This situation can be deduced by comparison of the sibling risk values for major loci in genome-wide linkage scans with the total genetic component estimated from twin and family studies. Although major loci, such as the HLA region in type I diabetes and ankylosing spondylitis, clearly exist, they account for less than half the genetic effect.[4] The diversity of the genetic loci that have an effect on immunologically related diseases might be a result of evolutionary selection pressures and frequency-dependent mechanisms that maintain polymorphic alleles that confer resistance to specific pathogens. Characterisation of more and more specific interactions between variable alleles of the host and pathogen supports this view. A major unaddressed issue is the extent of interactions between susceptibility loci. Such epistatic effects could be important in defining an individual's overall susceptibility profile, but current study designs have little power to detect such occurrences.

## HLA associations

A complete sequence of the MHC and precise definition of the diversity of many hundreds of HLA alleles now permits analysis of the role of this genetic region with unprecedented precision. Despite the impetus provided by the strong associations with ankylosing spondylitis and narcolepsy, these links are exceptions to the rule that, in general, associations between HLA and autoimmune and infectious diseases are modest, and the precise allelic

*Lancet* 2001; **357:** 2037–41

**Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK** (A V S Hill ??)
(e-mail: adrian.hill@well.ox.ac.uk)



Adrian Hill

**Three-dimensional representation of a peptide in the groove of an HLA molecule**

associations sometimes differ between populations.

Appreciation of the likely size of HLA effects on disease susceptibility, and the ever greater number of alleles to be defined, have necessitated new approaches to the design of such studies. The key difficulty is lack of statistical power resulting from the large number of alleles and haplotypes that might correlate with the disease. One strategy gaining recognition is to identify a possible association by typing all loci and variants in a proportion of the sample and to attempt to replicate the result in the remaining samples.[5] The increasing sophistication of genetic association studies has been led by the HLA field, but encouraged by the popularity of association studies with more and more HLA genes. However, refinements such as family-based association studies, replication studies, more sophisticated matching and statistical analyses, and, above all, adequate sample size are not yet sufficiently appreciated by investigators and editors, and inspection of the resulting apparent inconsistencies in published work can lead to unjustified nihilism.[6]

Improvements in molecular genotyping have allowed disease associations with HLA to be narrowed down to specific subtypes. For example, ankylosing spondylitis is associated with the HLA-B27 subtypes HLA-B*2705 and B*2704, but probably not with B*2706.[7] This subdivision could facilitate the identification of the key epitopes of relevance to pathogenesis. Molecular analysis has shown an association between coeliac disease and a specific heterodimer of HLA-DQ2, in which the two constituent chains of this HLA molecule can be encoded by alleles on the same chromosome or on opposite chromosomes of a homologous pair.[8]

Zinkernagel and Doherty's[9] discovery of MHC restriction of cellular immune responses allowed them to speculate in 1975 that this phenomenon could have facilitated the evolutionary maintenance of the genetic diversity of HLA types in two ways. First, particular HLA types might provide protection against specific infectious pathogens, particularly viruses. Although HLA associations with mycobacterial diseases and malaria are well known, evidence of protective HLA class I and II associations with viral infections caused by hepatitis B virus, hepatitis C virus, human lymphotropic virus 1, HIV, and dengue virus has also been provided.[10] The second proposal was that individuals heterozygous at an HLA locus might be more protected than homozygotes, presumably (in modern conceptualisation) through the ability to present more peptide epitopes to T lymphocytes. Again, in the past few years, this concept has been supported by reports of heterozygotes for HLA class II and I genes being resistant to persistent hepatitis B and progression of HIV disease, respectively.[11,12] Whether particular HLA types correlate with susceptibility or resistance to molecularly defined pathogen epitopes is yet to be discovered; however, preliminary data on malaria support this possibility.[13]

## Cytokines, chemokines, and receptors

Promoter variants of various cytokine genes have been investigated for disease associations. In large studies, some clear associations have emerged—eg, with variants in the tumour necrosis factor gene promoter and severe malaria. These links have been supported by functional studies such as electrophoretic mobility shift assays to identify possible interacting transcription factors, and transfection assays, usually in transformed cell lines.[14] Associations between the interleukin 1 gene cluster and several disorders, including Alzheimer's disease and inflammatory bowel disease, have been reported,[15–18] and an association with gastric cancer and *Helicobacter pylori*-induced hypochlorhydria has been supported by DNA–protein interaction studies of a promoter variant in the interleukin 1β gene.[19] Apparently functional variants of the interleukin 10 gene promoter have also been described, and there is preliminary evidence of associations with arthritis, ulcerative colitis, and viral infections.[20–23] Similarly, variants of interleukin 6 and interleukin 4 might be of relevance to juvenile chronic arthritis and atopy, repectively.[24,25]

In 1996, several groups (eg, ref 26) reported that a 32-bp deletion in the chemokine-receptor-5 (*CCR5*) gene was associated with resistance to HIV-1 infection and reduced

| Disease | Population | Locus | Allele/variant | Association |
|---|---|---|---|---|
| Ankylosing spondylitis[7] | Many | HLA-B | *2705 | Susceptibility |
| Coeliac disease[8] | Whites | HLA-DQ | DQ2 heterodimer | Susceptibility |
| Hepatitis C persistence[56] | Europeans | HLA-DRB1 | *1101 | Viral clearance |
| HIV infection[26] | Whites | CCR5 | 32 bp deletion | Resistance |
| Hypochlorhydria[19] | Whites | Interleukin 1 | Promoter | Susceptibility |
| Malaria[57] | Africans | HLA-B | *5301 | Resistance |
| Tuberculosis[31] | Africans | NRAMP1 | Various | Susceptibility |
| Type I diabetes[58] | Whites | Insulin | Minisatellite | Susceptibility |
| Type II diabetes[59] | Whites | Calpain 10 | Intronic variant | Susceptibility |

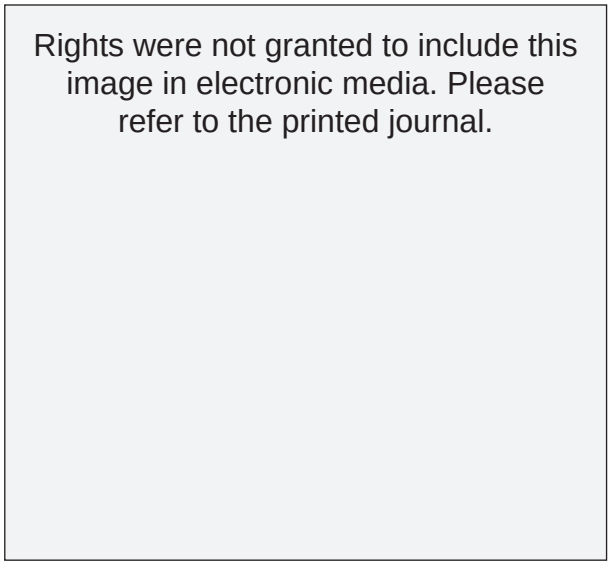
**Selected immunogenetic associations with disease**

rate of disease progression to AIDS. This strongly protective allele is absent in sub-Saharan Africans in whom HIV infection is most prevalent. Variation in the flanking gene for the CCR2 chemokine receptor is relevant to disease progression in some populations,[27] and a promoter variant of the gene for one of the ligands for CCR5—RANTES—has been associated with disease progression in Japanese people.[28] These genetic studies have stimulated efforts to design blockers of the CCR5 co-receptor that might interfere with HIV invasion and disease progression.[29] By comparison with the efforts made on cytokine promoters, much less is known of the relevance of polymorphism in the larger genes for cytokine receptors, an area which might prove as productive.

## Other candidate genes

Many other genes have now been analysed in detail to assess possible disease associations (table); I will mention only a selected few. Genetic linkage and positional cloning studies in mice identified the murine homologue of the natural-resistance-associated macrophage protein 1 (*NRAMP1*) gene as a major locus for susceptibility to some *Leishmania* parasites, and certain strains of *Salmonella* and *Mycobacterium bovis*.[30] In several populations, variants of the human *NRAMP1* gene have been associated with tuberculosis.[31] The effect is much smaller than in murine studies, but it has been reproducible. The same gene could be associated with sarcoidosis[32] and juvenile rheumatoid arthritis.[33] Variants of the vitamin D receptor, which are associated with osteoporosis in some populations, have also been implicated in tuberculosis.[34]

There is increasing interest in the links between innate and acquired immunity, and variants of the newly identified toll receptors could have a role in determining the type of immune response that develops.[35] Many researchers have studied the common deficiency alleles of mannose-binding lectin, which is potentially an important component of innate immunity through its ability to opsonise pathogens and activate complement. However, published claims of associations between mannose-binding lectin and susceptibility to various infectious diseases have been equivocal owing to issues of study design or genotyping, or to small sample size.[36] Functional variants of the Fcγ RII (or CD32) locus were suggested as relevant to disease caused by invasive encapsulated bacteria, and evidence of a protective effect of the H/H131 genotype against severe meningococcal disease is accumulating.[37] The platelet-membrane glycoprotein and scavenger receptor CD36 might have a role in lipid metabolism, host defence, and haemostasis. Inactivating and truncating mutations of this gene are common in African individuals and some other populations, and the potential role of CD36 in sequestration of malaria parasites and altered maturation of dendritic cells initially suggested that CD36 could be added to the list of genes selected by malaria parasites; recent data do not support this hypothesis, however.[38]




Science Photo Library

**Positive results on a DNA Biochip test**

In general, identification of loci, even in the MHC, that affect the level of measured immune responses has been more difficult than identification of those that affect susceptibility to disease. However, there is now preliminary evidence that non-HLA genes—eg, interleukin 4, cytotoxic T lymphocyte-associated molecule 4 (*CTLA4*), and *NRAMP1*—might be relevant.[39] CTLA4 expressed on T cells interacts with co-stimulatory molecules on antigen-presenting cells, and an aminoacid change in exon 1 might be relevant to type I diabetes and autoimmune thyroid disease.[40,41]

## Genome-wide searches

Mapping and analysis of the human genome has allowed the investigation of candidate susceptibility genes to be supplemented by a fundamentally different form of genetic analysis. Genome-wide linkage studies have become important in the analysis of all complex genetic disease. In such studies, the transmission of variants is followed within families, and analyses are done to detect increased co-segregation of disease with a few hundred informative microsatellite markers chosen to cover evenly the 23 chromosomes. This approach requires multicase families—typically affected sibling pairs—and substantial efforts have now been made to recruit these individuals for various complex diseases. Although linkage studies have been successful in identifying the genes for many rare Mendelian subphenotypes within complex disease, commoner variants in these genes seldom seem to be involved in general susceptibility.

Genome scans have now been reported for a large number of common complex diseases including type I and II diabetes, multiple sclerosis, atopy and asthma, rheumatoid arthritis, coeliac disease, inflammatory bowel disease, and tuberculosis. The emerging picture is one in which occasional major loci stand out, but in which most of the genetic component seems to be encoded by many minor polygenes. These polygenes are generally undetectable by linkage analysis in collections of fewer than 1000 families, and efforts are focusing on gene identification for the chromosomal regions that produce consistent evidence of a major effect in different populations. Successes include the apolipoprotein E gene in Alzheimer's disease,[42] the calpain-10 locus in type II diabetes,[43] and *NOD2* in Crohn's disease.[44–46] A region of chromosome 5q encoding numerous cytokines has been linked to asthma, schistosome egg burden, and malaria parasite density.[47–49] For most linkages, fine mapping and gene identification has been more time-consuming than expected, but such endeavours should be helped by the hundreds of thousands of new SNP markers becoming available through an international consortium.

But what of the polygenes that do not show up on genome-wide linkage analysis, and of diseases such as multiple sclerosis in which a substantial genetic component seems to be widely distributed without any major non-MHC loci? Again, technical advances promise a solution.

The possibility that genome-wide analysis could be undertaken by association rather than by linkage is attracting much attention. Many more markers would be needed to cover the genome, because genetic markers need to be much closer together to stay together in the population rather than in a family. But experts disagree on whether this number of markers would be 30 000 or 300 000.[50] The variety of alleles that can affect a gene's function is also an problem for this approach, because different mutations will commonly have different flanking markers, but with rapid increases in genotyping capacity, answers will doubtless be forthcoming.

A particular impetus for these genome-wide approaches remains the hope of finding the unexpected. Often much more can be learned from identifying a surprising gene in an unsuspected pathway than from proving an effect of a likely candidate gene.

## Immunogenomic applications

There are several reasons for the current upsurge in funding for genetic and genomic analyses of complex disease. The greatest potential dividend for pharmaceutical and biotechnology companies is the identification of new molecular targets for intervention with pharmaceutical and biological drugs. Immunogenetic variants could be particularly helpful in identifying indications for the large number of immunological mediators still seeking a therapeutic indication. Associations with variants in receptors and in intracellular transduction and signalling molecules might be of particular use for this goal. Also valuable would be the identification of a role for an unexpected or unidentified whole pathway that provides a series of new target molecules. With the completion of the genome sequence, the potential for uncovering new pathways involved in pathogenesis is huge, and although more obvious in poorly understood diseases such as autism, is still substantial in most complex diseases.

Another application of increasing interest is the prediction of the likely effectiveness of particular therapies according to the genotypes of individuals. Immunogenetic polymorphisms could be used to tailor the use of expensive, partly effective, immunotherapies such as interferon alfa in persistent hepatitis. Immunogenetic profiling of patients with a large array of common DNA variants is likely to rationalise the use of expensive biological treatments in the future.

Molecular analysis of well defined HLA associations are beginning to have an effect on the design of new vaccines and immunotherapeutics. Two research groups have identified what seems to be the key toxic HLA-DQ2-restricted epitope in gliadin,[51,52] and a range of epitope-based immunotherapeutic approaches for coeliac disease are being explored. HLA associations with resistance to infectious diseases have implicated certain T-cell immune responses in protective immunity, and epitope-based vaccines[53] based on these findings are now in clinical trials.

## Prospects

The availability of many more polymorphisms in immunologically related and other genes will increase further the power and the popularity of genetic analyses of complex diseases. The rate of progress will partly be determined by the availability of new high-throughput genotyping technologies. The use of "gene chips" or microarrays to determine genotypes is less developed than for expression profiling, and, although already available for some loci,[54] whether hybridisation-based or primer extension chips will emerge as the most useful rapid typing approach is unclear. Competing technologies include mass spectrometry, flow-based methods with luminescent beads, various fluorescence-based DNA methods such as Invader (Third Wave Technologies, Madison, WI, USA), and SNiPer (Amersham Pharmacia Biotech, UK) assays. Also, methods of estimation of allele frequencies in pooled samples will probably contribute to improved high-throughput genotyping strategies.

One effect on study design will be the capacity to analyse multiple polymorphisms in any gene of interest and in nearby flanking genes. This ability should help address whether the associated gene is causally involved or merely a marker. As genotyping capacity increases, study design will become more important. Published studies on immunogenetics contain too many underpowered studies of just a few markers in sample sizes of only a hundred or so cases, even for common diseases. Many such associations are not repeatable and editors are often reluctant to publish larger negative studies.[55] Approaches for dealing with such preliminary priority claims on small samples are needed, as is making available an increasing body of useful but negative findings with candidate genes. Appropriately monitored websites could address both needs.

Another suggested solution to the difficulty of interpreting genetic association data is to require evidence that any associated polymorphisms are of functional significance. But this step is, arguably, retrograde. The greatest use for genetics lies in uncovering the unknown, not in validating functional assays often undertaken in artificial in-vitro conditions. For examples, expression assays of cytokine promoter variants are often done in transformed cells of a particular lineage in rather artificial microenvironments. A convincing replicated association with non-functional polymorphisms in a large well designed study might be more valuable than a tenative association with a variant of probable functional significance.

There is no doubt that genomics is going to affect all areas of clinical research, and design issues that were once of concern mainly to the HLA-typing community are now receiving much wider consideration. Genetic variants of any immunogically relevant gene showing a suggestive disease association are rapidly screened in many other diseases. Immunogenetics has flown free of the MHC and embraced a new era of immunogenomics. This integrative force in medicine driven by the huge potential of genomics is a welcome counter to increasing clinical specialisation.

### References


1 Risch NJ. Searching for genetic determinants in the new millennium. *Nature* 2000; **405:** 847–56.

2 Jepson A, Banya W, Sisay-Joof F, 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* 1997; **65:** 872–76.

3 Risch N. Assessing the role of HLA-linked and unlinked determinants of disease. *Am J Hum Genet* 1987; **40:** 1–14.

4 Brown MA, Kennedy LG, MacGregor AJ, et al. Susceptibility to ankylosing spondylitis in twins: the role of genes, HLA, and the environment. *Arthritis Rheum* 1997; **40:** 1823–28.

5 Thursz M, Yallop R, Goldin R, Trepo C, Thomas HC. Influence of MHC class II genotype on outcome of infection with hepatitis C virus. *Lancet* 1999; **354:** 2119–24.

6 Gambaro G, Anglani F, D'Angelo A. Association studies of genetic polymorphisms and complex disease. *Lancet* 2000; **355:** 308–11.

7 Lopez-Larrea C, Sujirachato K, Mehra NK, et al. HLA-B27 subtypes in Asian patients with ankylosing spondylitis: evidence for new associations. *Tissue Antigens* 1995; **45:** 169–76.

8 Sollid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E. Evidence for a primary association of celiac disease to a particular HLA-DQ alpha/beta heterodimer. *J Exp Med* 1989; **169:** 345–50.

9 Doherty PC, Zinkernagel RM. A biological role for the major histocompatibility antigens. *Lancet* 1975; **1:** 1406–09.

10 Hill AV. The immunogenetics of human infectious diseases. *Annu Rev Immunol* 1998; **16:** 593–617.

11 Thursz MR, Thomas HC, Greenwood BM, Hill AV. Heterozygote advantage for HLA class-II type in hepatitis B virus infection. *Nat Genet* 1997; **17:** 11–12.

12 Carrington M, Nelson GW, Martin MP, et al. HLA and HIV-1: heterozygote advantage and B*35-Cw*04 disadvantage. *Science* 1999; **283:** 1748–52.

13 Gilbert SC, Plebanski M, Gupta S, et al. Association of malaria parasite population structure, HLA, and immunological antagonism. *Science* 1998; **279:** 1173–77.

14 McGuire W, Knight JC, Hill AV, Allsopp CE, Greenwood BM, Kwiatkowski D. Severe malarial anemia and cerebral malaria are associated with different tumor necrosis factor promoter alleles. *J Infect Dis* 1999; **179:** 287–90.

15 Nicoll JA, Mrak RE, Graham DI, et al. Association of interleukin-1 gene polymorphisms with Alzheimer's disease. *Ann Neurol* 2000; **47:** 365–68.

16 Grimaldi LM, Casadei VM, Ferri C, et al. Association of early-onset Alzheimer's disease with an interleukin-1alpha gene polymorphism. *Ann Neurol* 2000; **47:** 361–65.

17 Du Y, Dodel RC, Eastwood BJ, et al. Association of an interleukin 1 alpha polymorphism with Alzheimer's disease. *Neurology* 2000; **55:** 480–83.

18 Nemetz A, Nosti-Escanilla MP, Molnar T, et al. IL1B gene polymorphisms influence the course and severity of inflammatory bowel disease. *Immunogenetics* 1999; **49:** 527–31.

19 El-Omar EM, Carrington M, Chow WH, et al. Interleukin-1 polymorphisms associated with increased risk of gastric cancer. *Nature* 2000; **404:** 398–402.

20 Eskdale J, McNicholl J, Wordsworth P, et al. Interleukin-10 microsatellite polymorphisms and IL-10 locus alleles in rheumatoid arthritis susceptibility. *Lancet* 1998; **352:** 1282–83.

21 Tagore A, Gonsalkorale WM, Pravica V, et al. Interleukin-10 (IL-10) genotypes in inflammatory bowel disease. *Tissue Antigens* 1999; **54:** 386–90.

22 Helminen M, Lahdenpohja N, Hurme M. Polymorphism of the interleukin-10 gene is associated with susceptibility to Epstein-Barr virus infection. *J Infect Dis* 1999; **180:** 496–99.

23 Edwards-Smith CJ, Jonsson JR, Purdie DM, Bansal A, Shorthouse C, Powell EE. Interleukin-10 promoter polymorphism predicts initial response of chronic hepatitis C to interferon alfa. *Hepatology* 1999; **30:** 526–30.

24 Fishman D, Faulds G, Jeffery R, et al. The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis. *J Clin Invest* 1998; **102:** 1369–76.

25 Walley AJ, Cookson WO. Investigation of an interleukin-4 promoter polymorphism for associations with asthma and atopy. *J Med Genet* 1996; **33:** 689–92.

26 Liu R, Paxton WA, Choe S, et al. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell* 1996; **86:** 367–77.

27 Smith MW, Dean M, Carrington M, 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* 1997; **277:** 959–65.

28 Liu H, Chao D, Nakayama EE, et al. Polymorphism in RANTES chemokine promoter affects HIV-1 disease progression. *Proc Natl Acad Sci USA* 1999; **96:** 4581–85.

29 Proudfoot AE, Wells TN, Clapham PR. Chemokine receptors: future therapeutic targets for HIV? *Biochem Pharmacol* 1999; **57:** 451–63.

30 Vidal SM, Malo D, Vogan K, Skamene E, Gros P. Natural resistance to infection with intracellular parasites: isolation of a candidate for Bcg. *Cell* 1993; **73:** 469–85.

31 Bellamy R, Ruwende C, Corrah T, McAdam KP, Whittle HC, Hill AV. Variations in the NRAMP1 gene and susceptibility to tuberculosis in West Africans. *N Engl J Med* 1998; **338:** 640–44.

32 Maliarik MJ, Chen KM, Sheffer RG, et al. The natural resistance-associated macrophage protein gene in African Americans with sarcoidosis. *Am J Respir Cell Mol Biol* 2000; **22:** 672–75.

33 Sanjeevi CB, Miller EN, Dabadghao P, et al. Polymorphism at NRAMP1 and D2S1471 loci associated with juvenile rheumatoid arthritis. *Arthritis Rheum* 2000; **43:** 1397–404.

34 Bellamy R. Evidence of gene-environment interaction in development of tuberculosis *Lancet* 2000; **355:** 588–89.

35 Arbour NC, Lorenz E, Schutte BC, et al. TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. *Nat Genet* 2000; **25:** 187–91.

36 Tang C, Kwiatkowski D. Mannose-binding lectin and meningococcal disease. *Lancet* 1999; **354:** 336–37.

37 Platonov AE, Shipulin GA, Vershinina IV, Dankert J, van de Winkel JG, Kuijper EJ. Association of human Fc gamma RIIa (CD32) polymorphism with susceptibility to and severity of meningococcal disease. *Clin Infect Dis* 1998; **27:** 746–50.

38 Aitman TJ, Cooper LD, Norsworthy PJ, et al. Malaria susceptibility and CD36 mutation. *Nature* 2000; **405:** 1015–16.

39 Blackwell JM, Black GF, Sharples C, Soo SS, Peacock CS, Miller N. Roles of Nramp1, HLA, and a gene(s) in allelic association with IL-4, in determining T helper subset differentiation. *Microbes Infect* 1999; **1:** 95–102.

40 Marron MP, Raffel LJ, Garchon HJ, et al. Insulin-dependent diabetes mellitus (IDDM) is associated with CTLA4 polymorphisms in multiple ethnic groups. *Hum Mol Genet* 1997; **6:** 1275–82.

41 Donner H, Rau H, Walfish PG, et al. CTLA4 alanine-17 confers genetic susceptibility to Graves' disease and to type 1 diabetes mellitus. *J Clin Endocrinol Metab* 1997; **82:** 143–46.

42 Roses AD. A model for susceptibility polymorphisms for complex diseases: apolipoprotein E and Alzheimer disease. *Neurogenetics* 1997; **1:** 3–11.

43 Horikawa Y, Oda N, Cox NJ, et al. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nat Genet* 2000; **26:** 163–75.

44 Hugot HP, Chamaillard M, Zouali H, et al. Association of *NOD2* leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001; **411:** 599–603.

45 Ogura Y, Bonen DK, Inohara N, et al. A frameshift mutation in *NOD2* associated with susceptibility to Crohn's disease. *Nature* 2001; **411:** 603–06.

46 Hampe J, Cuthbert A, Croucher PJP, et al. Association between insertion mutation in *NOD2* gene and Crohn's disease in British and German populations. *Lancet* 2001; **357:** 1925–28.

47 Marquet S, Abel L, Hillaire D, et al. Genetic localization of a locus controlling the intensity of infection by *Schistosoma mansoni* on chromosome 5q31-q33. *Nat Genet* 1996; **14:** 181–84.

48 Rihet P, Traore Y, Abel L, Aucan C, Traore-Leroux T, Fumoux F. Malaria in humans: *Plasmodium falciparum* blood infection levels are linked to chromosome 5q31-q33. *Am J Hum Genet* 1998; **63:** 498–505.

49 Lonjou C, Barnes K, Chen H, et al. A first trial of retrospective collaboration for positional cloning in complex inheritance: assay of the cytokine region on chromosome 5 by the consortium on asthma genetics (COAG). *Proc Natl Acad Sci USA* 2000; **97:** 10942–47.

50 Kruglyak L. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet* 1999; **22:** 139–44.

51 Anderson RP, Degano P, Godkin AJ, Jewell DP, Hill AV. In vivo antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. *Nat Med* 2000; **6:** 337–42.

52 Arentz-Hansen H, Korner R, Molberg O, et al. The intestinal T cell response to alpha-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. *J Exp Med* 2000; **191:** 603–12.

53 Gilbert SC, Plebanski M, Harris SJ, et al. A protein particle vaccine containing multiple malaria epitopes. *Nat Biotechnol* 1997; **15:** 1280–84.

54 Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ. High density synthetic oligonucleotide arrays. *Nat Genet* 1999; **21** (Suppl 1)**:** 20–24.

55 Keavney B, McKenzie C, Parish S, et al. Large-scale test of hypothesised associations between the angiotensin-converting-enzyme insertion/deletion polymorphism and myocardial infarction in about 5000 cases and 6000 controls. International Studies of Infarct Survival (ISIS) Collaborators. *Lancet* 2000; **355:** 434–42.

56 Thursz M, Yallop R, Goldin R, Trepo C, Thomas HC. Influence of MHC class II genotype on outcome of infection with hepatitis C virus. *Lancet* 1999; **354:** 2119–24.

57 Hill AVS, Allsopp CE, Kwiatkowski D, et al. Common west African HLA antigens are associated with protection from severe malaria. *Nature* 1991; **352**: 595–600.

58 Bell GI, Horita S, Karam JH. A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. *Diabetes* 1984; **33**: 176–83.

59 Horikawa Y, Oda N, Cox NJ, et al. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nat Genet* 2000; **26:** 163–75.