



The genetics of complex autoimmune diseases: non-MHC susceptibility genes

Amy Wandstrat and Edward Wakeland

Susceptibility to complex autoimmune diseases (AIDs) is a multigenic phenotype affected by a variety of genetic and environmental or stochastic factors. After over a decade of linkage analyses, the identification of non-major histocompatibility complex (non-MHC) susceptibility alleles has proved to be difficult, predominantly because of extensive genetic heterogeneity and possible epistatic interactions among the multiple genes required for disease development. Despite these difficulties, progress has been made in elucidating the genetic mechanisms that influence the inheritance of susceptibility, and the pace of gene discovery is accelerating. An intriguing new finding has been the colocalization of several AID susceptibility genes in both rodent models and human linkage studies. This may indicate that several susceptibility alleles affect multiple AIDs, or alternatively that genomic organization has resulted in the clustering of many immune system genes. The completion of the human genome sequence, coupled with the imminent completion of the mouse genome, should yield key information that will dramatically enhance the rate of gene discovery in complex conditions such as AID susceptibility.

Complex autoimmune diseases (AIDs) are chronic conditions initiated by a loss of immunologic tolerance to self-antigens. Clinical disease generally manifests as a result of damage induced in one or more organ systems *via* the inappropriate activation of immune-mediated inflammation. Collectively, AID is estimated to affect 4–5% of the population, females generally having a higher disease incidence than males¹. Six of the most common AIDs are rheumatoid arthritis (RA), Graves' disease, insulin-dependent diabetes mellitus or type I (autoimmune) diabetes (IMD) pernicious anemia, systemic lupus erythematosus (SLE) and multiple sclerosis (MS); collectively they represent about 50% of all AIDs. Roughly 1 in 30 individuals is afflicted with some type of autoimmune disease, thus making autoimmunity a major health problem in modern medicine.

Although the pathogenic mechanisms responsible for the initiation of autoimmunity remain poorly understood, a variety of classic studies have clearly demonstrated that genetic predisposition is a major factor

in disease susceptibility. The most potent genetic influence on susceptibility to autoimmunity is the major histocompatibility complex (MHC), which has been known for over two decades to affect susceptibility to a variety of AIDs. We will focus here on non-MHC susceptibility genes; for a review of MHC genes and autoimmunity see².

Information obtained *via* linkage studies of AIDs in both humans and rodents has begun to elucidate the role of genetics in disease predisposition. Here, we will provide an overview of the genetic mechanisms affecting the inheritance of susceptibility and discuss the progress that has been made toward identifying non-MHC genes and genetic pathways that contribute to AID susceptibility. The identification of these non-MHC susceptibility genes is expected to provide insights into the mechanisms that mediate disease pathogenesis and possibly identify new targets for the development of therapeutic strategies.

Genetic predisposition to AIDs

The powerful impact of genetic predisposition on susceptibility to autoimmunity was first identified by the analysis of disease concordance rates in monozygotic twins. The monozygotic disease concordance rate ranges from about 15% for RA³ to a fairly robust 57% for SLE⁴ (Table 1). Comparisons of these high concordance rates with disease incidence in the general population predict that genetic predisposition is the dominant factor in AID susceptibility. The dramatic decrease in the concordance rate of siblings compared with that of monozygotic twins supports the presence of multiple genes contributing to the genetic predisposition. Finally, the calculation of λ s for these diseases, which are the ratio of the risk of disease recurrence among the siblings of affected individuals to disease incidence in the general population, also supports a potent role for genetic predisposition in disease susceptibility. Table 1 shows λ s for various AIDs, which range from about 10 for RA to as high as 40 for SLE⁴.

The powerful influence of genetic predisposition on AID susceptibility was initially interpreted as indicating that genome-wide linkage analysis would allow the identification of many potent non-MHC AID susceptibility genes. This expectation fostered the development of international coalitions focused on collecting large cohorts of multiplex families afflicted with specific AID and utilizing state-of-the-art technologies to scan their genomes for the locations of susceptibility genes^{5–15}. After over a decade of such analyses, the inheritance of AID susceptibility has proved to be highly complex and not readily amenable to genetic analysis in heterogeneous populations.

The consistent observation throughout these genome scans of AIDs has been the detection of many genomic segments exhibiting a weak statistical association with disease susceptibility. Individual genomic intervals are in general associated with susceptibility to AID, with lod scores ranging from 2.0 to occasionally approaching 5.0^{5–15}. For comparison, a fully penetrant Mendelian disease locus would be detected with a lod score approaching 30 by the analysis of 100 affected sibling pairs, which



would be a small sample size for most characterizations of AID susceptibility genetics. In addition, mapping studies by separate investigators working on the same AID frequently do not detect an association to the same genomic regions, thus raising an issue of reproducibility for many of the associations reported. The prevailing situation in most AIDs is therefore that susceptibility is only modestly associated with any specific non-MHC locus, despite the potent role for genetic predisposition in disease susceptibility.

The resolution of this apparent paradox lies in the complexity of AID genetics. The inheritance of AID susceptibility is multifactorial, which means that susceptibility arises from the combined impact of multiple contributing susceptibility genes, each potentially interacting with a poorly defined array of environmental and/or stochastic factors. In such a complex system, sample size rapidly limits the feasibility of obtaining statistically significant associations. In addition, the detection of these loci is complicated by two factors that commonly influence the inheritance of multifactorial traits: genetic heterogeneity and epistasis.

Genetic heterogeneity

Genetic heterogeneity refers to the presence of multiple combinations of genes within the genome that are capable of causing a similar or identical disease phenotype. Genetic heterogeneity is a common feature of many genetic systems in both humans and animal models. It simply reflects the fact that many genes participate in the development of complex phenotypes and that different combinations of genetic abnormalities can lead to a similar outcome.

Indications of a significant degree of genetic heterogeneity in AID susceptibility are apparent from linkage studies in both human populations and animal models. Classic association studies for candidate susceptibility genes in both IMD and SLE have, for example, frequently observed significant variations between ethnic groups in the disease association of specific alleles and disease phenotypes^{16,17}. Linkage analyses have detected an increase in statistical associations with specific genomic intervals only in specific ethnic groups^{6,18}. These results are readily explained by variations in the genetic basis for predisposition to AID between ethnic groups, although definitive results must await the identification and analysis of specific susceptibility alleles.

Comparisons of the genomic locations of susceptibility genes in separate mouse models of IMD, SLE, experimental allergic encephalomyelitis (EAE, an animal model of MS) and collagen-induced RA also clearly indicate that the genomic locations of many susceptibility alleles vary between models. Although attention has generally focused on the colocalization of susceptibility genes among AID-prone strains (see below), most of the genomic segments detected are not shared between different animal models, even with the same AID^{19–26}. A recent linkage analysis of the lupus-prone BXSB strain found, for example, that only two of five intervals overlapped with intervals detected in linkage studies performed in other strain combinations. This indicated that lupus susceptibility was being mediated in BXSB by loci that were not involved in disease susceptibility in NZM2410, NZB/NZW or MRL/*lpr* strains²⁶. These results are representative of findings in other AID models and indicate that susceptibility is mediated predominantly by a heterogeneous array of genes in murine models of AID.

Epistatic interactions

Epistasis is classically defined as a genetic interaction in which the genotype at one locus affects the phenotypic expression of the genotype

Table 1. Relative risks in autoimmune disease

Disease	Concordance rate (%)			Population prevalence (%)	λ_s
	Monozygotic twins	Dizygotic twins	Non-twin siblings		
IMD	30–50	0–13	6	0.4	15 ^a
MS	25	0–5	3–5	0.1	20 ^b
SLE	24–57	2–5	2–5	0.2	20–40 ^c
RA	12–15	3–4	2–4	0.24–1.0	5–10 ^d

^aRefs. 106 and 107. ^bRef. 49. ^cRef. 108. ^dRef. 3.

at another locus. Evidence consistent with epistatic interactions among susceptibility alleles has been reported in both animal models and human linkage studies of AID^{27–30}. Synergism between susceptibility alleles is, for example, clearly seen in a recent analysis of the congenic strains B6.*Sle1*, B6.*yaa*, and B6.*Sle1/yaa*. B6.*Sle1* and B6.*yaa* are B6-congenic mice that carry the *Sle1* and *yaa* susceptibility genes for systemic autoimmunity, respectively. Each strain spontaneously produces nonpathogenic autoantibodies to nuclear antigens but fails to develop severe autoimmunity, having a normal lifespan and developing little or no glomerulonephritis. However, when these two susceptibility alleles are combined in the B6.*Sle1/yaa* bicongenic strain, a severe systemic autoimmunity develops, which culminates in fatal glomerulonephritis with an incidence of 70% by 9 months of age³⁰. This is an example of epistasis between two susceptibility alleles leading to a greater increase in disease severity than would be predicted by simply adding together their individual phenotypes.

A second type of epistasis, in which the autoimmune phenotypes of susceptibility alleles are suppressed by epistatic modifiers, has also been detected in an animal model of SLE³¹. Suppressive modifiers were detected *via* the analysis of the disease phenotype mediated by *Sle1*, *Sle2* and *Sle3* when introgressed onto different genetic backgrounds. These three genes (or gene clusters, see below) are the primary genes responsible for susceptibility to lupus nephritis in the NZM2410 lupus-prone mouse³². When they are introgressed from NZM2410 onto the B6 background, the resultant B6.*Sle1/Sle2/Sle3* triple congenic strain develops fatal lupus nephritis with a penetrance approaching 100% in both genders at 9 months of age. Although all three of these susceptibility alleles are derived from the NZW genome, NZW exhibits very benign autoimmune phenotypes that develop only in females older than 12 months³³. Thus, the expression of *Sle1*, *Sle2* and *Sle3* is significantly suppressed in NZW mice.

A linkage analysis designed to detect these suppressive modifiers in the NZW genome found four separate loci that accounted for the suppression of lupus susceptibility in NZW mice³¹. These results indicate that the disease mediated by susceptibility genes can be fully suppressed by other “modifying” genes in the genome, resulting in the development of a relatively normal immune phenotype, despite the presence of highly potent autoimmune disease alleles. The presence of similar suppressive modifiers in AID in humans has not been demonstrated, although it is reasonable to predict that similar genetic interactions will affect disease predisposition in humans.

A review of the extensive linkage studies that have accumulated in animal models of AID over the past 10 years suggests that epistatic interactions may be a common element in the complex inheritance of disease susceptibility. Some susceptibility alleles are, for example, inherited from the genome of the “normal” parental strain rather than the autoimmune-prone strain, suggesting that normal strains often

carry alleles that enhance disease susceptibility when integrated into a permissive genome^{32,34,35}. Similarly, several targeted gene disruptions have been reported to mediate autoimmune phenotypes, but only when crossed into a specific inbred strain or carried on a specific, mixed background³⁶⁻³⁸. Earlier analogous findings demonstrated that spontaneous mutations of Fas and Fas ligand lost most or all of their autoimmune phenotypes when crossed onto specific genetic backgrounds^{39,40}. These phenotypic variations likely represent epistatic interactions between susceptibility alleles and other unknown loci present in the genomes of various inbred mouse strains, although proof of this will require further linkage studies and possibly congenic strain construction.

Thus, epistasis appears to significantly affect the inheritance of susceptibility to AID. The data generated in animal models suggest that the potency of many susceptibility alleles is strongly dependent upon genomic context, as a result of complex interactions with other susceptibility alleles and suppressive modifiers. Although the extent of epistasis in human AID remains to be determined, it is reasonable to predict that epistatic interactions are a consequence of the many functional polymorphisms influencing immune recognition and responsiveness, and will therefore be a component of AID susceptibility in most species.

Environment, stochastic events and disease penetrance

Although genetic predisposition is the major factor dictating AID susceptibility, various data indicate that environmental or stochastic factors are also involved⁴¹⁻⁴³. A role for nongenetic factors in the initiation of disease was classically predicted by the incomplete concordance of disease expression among monozygotic twin pairs. Although concordance can be high in AID, it does not approach 100% (Table 1).

Thus, genetically predisposed individuals may or may not develop autoimmune disease, contingent upon other elements affecting their health. A similar interpretation can be drawn from the incomplete penetrance of spontaneous disease in inbred strains prone to the development of diabetes or lupus^{32,44-46}.

The strongest data supporting a role for environmental factors come from epidemiologic studies of MS⁴⁷⁻⁴⁹. Interestingly, the incidence of MS is distributed in a nonrandom fashion geographically, individuals at higher latitudes being at greater risk. In addition, some studies have demonstrated a higher disease incidence in certain locations, suggesting that disease risk is increased significantly by local environment. The biomedical basis for these observations is unknown, although some data have implicated microbial infections. Viral infections have also been implicated as environmental triggers in several other AID. Susceptibility to SLE has been, for example, associated with positive

seroconversion for Epstein-Barr virus (EBV) antibodies in a retrospective study of a large cohort of patients and matched controls^{50,51}. The strong impact of EBV infection on B cell activation makes this association intriguing, although more direct evidence will be required to establish a clear link between EBV and SLE.

Spontaneous disease incidence in AID-prone inbred strains would appear to be an excellent model in which to test the potential role of environmental triggers in the initiation of AID. In general, inbred strains of rodents that are susceptible to spontaneous AID, such as nonobese diabetic (NOD), MRL/*lpr*, BXSB and NZM2410, all exhibit an incomplete penetrance of disease. Because the environment in which mice are housed is highly controlled, including environments that are specific pathogen free (SPF), this argues that environmental variables are not necessary for the observed incomplete disease penetrance. The

incidence of diabetes dramatically increases in colonies of NOD mice housed under SPF conditions rather than conventional caging, indicating that risk increases in this inbred strain when microbial infection is limited⁵².

The incomplete penetrance of disease in AID-prone animal models has led to the hypothesis that disease can be initiated in highly susceptible animals by stochastic events that occur during the normal functioning of the immune system⁵³. In this regard, many multifactorial phenotypes exhibit incomplete penetrance that is similar in nature to that of AID in spontaneous models, this being commonly attributed to chance or stochastic events that occur during development or normal physiology⁵⁴. In the context of autoimmunity, a reasonable prediction is that genetic predisposition will lead to the creation of an immune system that contains significant autoimmune potential, but that disease is initiated only when specific T and/or B cells interact with a specific

self-antigen being presented appropriately in a stimulatory microenvironment. The probability of this “stochastic event” occurring will then dictate the relative risk for genetically identical individuals in a controlled environment.

Models of inheritance of disease susceptibility

The inheritance of multifactorial traits such as AID susceptibility is a complex process. Multifactorial inheritance was first described and modeled as a “threshold liability” in an analysis of polydactyly in guinea pigs⁵⁵. It was proposed that the penetrance of polygenic, qualitative phenotypes would increase in relation to the number of susceptibility genes present in the genome of an individual.

A hypothetical model of the inheritance of AID can be proposed (Fig. 1). The *x* axis of this graph defines increasing disease liability, the *y* axis represents the “threshold”, which delineates the point at

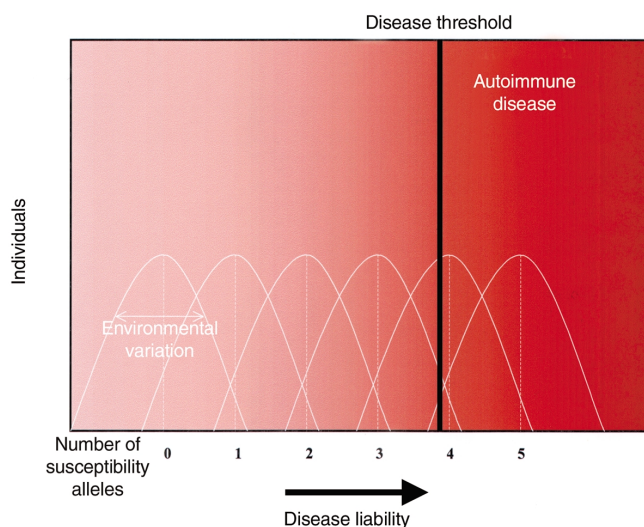


Figure 1. Threshold liabilities in autoimmune disease. In this model, only individuals located to the right of the disease threshold line will develop disease. The *x* axis represents increasing liability to disease, individuals being located on the *x* axis based on the degree of their predisposition to disease. An incremental increase in the number of susceptibility alleles progressively increases liability to disease, resulting in movement toward the disease threshold at the right on the *x* axis. The disease liability introduced by environmental and stochastic effects is represented by the normal distribution curve around the location of individuals with specific degrees of genetic predisposition for disease.



which individuals will develop disease. Genetic predisposition places individuals at some point along the x axis, based on the degree of susceptibility dictated by their genomes. Environmental and stochastic events will then increase or decrease their liability, depending on their life experience. These environmental factors are arbitrarily depicted as a normal distribution of liability around the mean location dictated by genetic predisposition. The exact distribution could potentially take any shape.

The inheritance of susceptibility would then be determined by the cumulative content of disease susceptibility that an individual inherits. The relationship presented in **Fig. 1** shows a simplified example of additive inheritance, in which each additional susceptibility allele that an individual inherits results in an incremental movement toward the disease threshold on the liability axis. The genomes of disease-prone inbred mouse strains such as NOD or NZM2410 contain enough susceptibility genes that their genetic predisposition places them beyond the disease threshold (**Fig. 1**). As a result, only a small fraction of individual mice in these strains will fail to develop disease, depending on the stochastic processes that occur during the normal functioning of their immune systems. Consistent with this model, AID-prone inbred strains contain many susceptibility genes, and disease penetrance decreases when the number of susceptibility genes is decreased *via* congenic strain construction⁴⁴.

The inheritance of disease susceptibility (**Fig. 1**) is presented in an extremely simplistic additive fashion; each additional susceptibility allele incrementally moves an individual an equivalent distance further toward the disease threshold. In reality, the process of inheritance is much more complex. As discussed above, epistatic interactions would be predicted to modify the incremental movement of individuals along the axis in a complex fashion that would not be additive. Thus, the position of an individual along the liability axis would be dependent upon the interactive consequences of all the susceptibility and suppressive alleles present in his or her genome.

In this regard, attempts to model the inheritance of AID susceptibility have often focused on distinguishing “additive” inheritance from “multiplicative” models⁵⁶. Linkage analyses in test crosses of AID-prone inbred strains have consistently found that relative risk increases in proportion to the number of active susceptibility genes present in the genome^{29,32,57,58}, but the goodness of fit for additive *versus* multiplicative models has not been established⁵⁹. Given the extensive genetic heterogeneity observed in AID inheritance, it is reasonable to predict that both models will be in some circumstances correct.

Future prospects for linkage analysis

Susceptibility to AID is among the most complex genetic systems currently being investigated. Several strategies to cope with difficulties resulting from genetic heterogeneity and epistasis are currently being employed by investigators. The subdivision of patient populations into more homogeneous groups based on an analysis of closely associated component phenotypes (such as the presence of specific autoantibodies or clinical features within a specific disease) may limit some elements of genetic heterogeneity. An alternative approach has been to focus on disease inheritance in isolated human subpopulations with limited ethnic heterogeneity, a strategy that was successfully employed for the identification of *AIRE* as the gene responsible for autoimmune polyglandular syndrome (type 1)^{60,61}. Although there has been some success in the more complex analysis of multifactorial AID susceptibility in isolated Scandinavian populations¹⁸, the value of this approach for the analysis of AID is controversial⁶². Overall, a

key element in the ultimate success of linkage analysis in AID will be the continued expansion of well characterized AID patient populations and families.

The recent publication of an almost complete nucleotide sequence of the human genome has provided detailed physical and molecular maps of the majority of human linkage groups⁶³. In addition, a growing database of single-nucleotide polymorphisms (SNPs), together with an improving technology for their detection, will greatly enhance the statistical power of linkage analysis^{64,65}. It is likely that a database identifying several SNPs for every gene in the human genome will be available within the next 2 or 3 years. Armed with these resources, many scientists are hoping to advance our knowledge of complex polygenic disease etiology by quickly identifying the genes involved in such traits.

Although the analysis of SNPs will undoubtedly provide new insights, the accuracy and sensitivity of association studies will nonetheless continue to be impeded by epistasis, incomplete penetrance and genetic and phenotypic heterogeneity. An additional factor affecting the accuracy of SNP analysis will be the poorly defined nature of genetic disequilibrium at the genomic level in human populations. That is, for strong linkage associations to be achieved, a SNP must be so tightly associated with a disease-causing allele that it can be used as a marker to identify the disease allele among individuals in outbred populations. The degree of linkage disequilibrium in populations for closely linked markers is now a major research focus in genetics. Current data suggest that the distance of tight association appears to be highly variable and that disequilibrium may persist over regions of less than 1 kb in certain parts of the genome. As a result, thousands of SNPs are likely to be required to scan the entire genome, which will lead to a significant increase in the cost of genetic analysis and diminish the statistical power for the detection of associations. Despite these difficulties and technical issues, the analysis of SNPs should dramatically improve the accuracy and power of linkage analysis in human populations.

Identifying disease genes

The ultimate goal of linkage analysis is the identification of genes and genetic pathways that mediate disease susceptibility. Once all the genes in the human genome have been identified and encoded in a database, the process of susceptibility gene identification will simplify to the determination of a precise genomic location. For monogenic diseases, linkage analysis often defines the genomic location of disease alleles with sufficient precision to limit the number of positional candidate genes to about a dozen. As discussed above, however, linkage analysis of AID has not yielded robust statistical correlations with any specific loci, so the precision with which susceptibility alleles are positioned is extremely poor, resulting in genomic segments that contain hundreds or even thousands of positional candidate genes. This has complicated the final stage of disease gene identification and has been the major deterrent to the identification of human susceptibility alleles. Transmission disequilibrium testing (TDT) has been the best strategy for narrowing the interval size and potentially identifying disease genes in human genetics⁶⁶. This strategy has been used to identify and exclude several strong candidates in AID^{67–69}.

Investigators working on human AIDs have accumulated evidence supporting the identification of a handful of susceptibility alleles. The first non-MHC susceptibility allele to be identified was a variable number tandem repeat polymorphism in the promoter of the insulin gene that affects susceptibility to diabetes^{70,71}. Several studies have clearly delineated multiple functional consequences of this polymorphism on the expression of both insulin and a closely linked insulin-like growth factor gene^{72–74}.

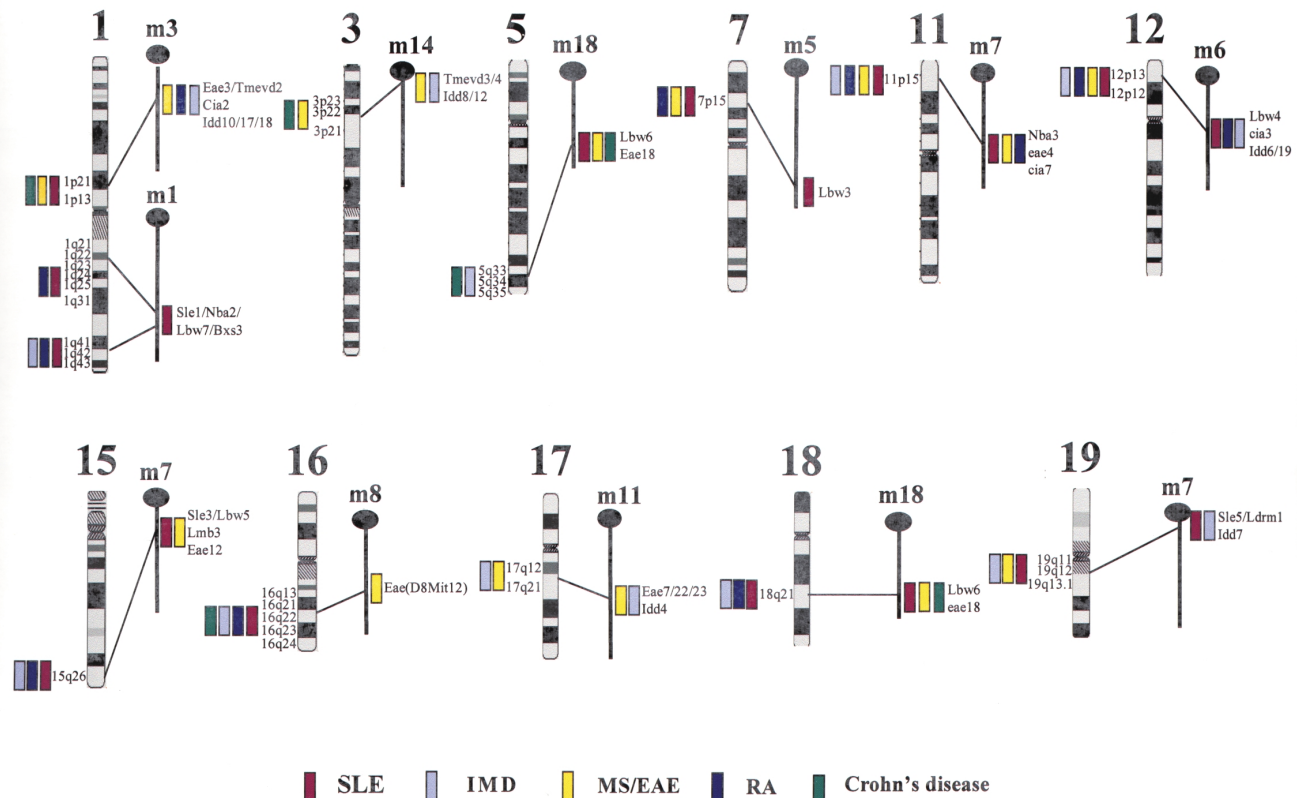


Figure 2. Chromosomal susceptibility regions associated with autoimmune disorders detected in human and mouse genome scans. Susceptibility regions are shown at their approximate positions^{96,97,109}. Quantitative trait loci (QTLs) for mouse models of SLE are designated *Sle*, *Lbw*, *Nba* (NZB/NZW), *Ldrn* (MRL/lpr) and *Bxs* (BXS). QTLs in MS mouse models are *Eae* and *Tmevd* (for Theiler's murine encephalomyelitis virus-induced demyelinating disease) and for IMD are *Idd*. QTLs for the RA mouse model are *Cia* (for collagen-induced arthritis). QTLs locations for Crohn's disease were based on a study of mouse inflammatory bowel disease model induced by dextran sulfate sodium⁴⁷. QTL locations are based on linkage maps available at the Mouse Genome Database (<http://www.informatics.jax.org/>). Relative chromosomal sizes and syntenic relationships are based on human-mouse homology maps available at the National Center for Biotechnology website (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Genome>).

The exact mechanism by which these expression variations mediate susceptibility to IMD remains to be elucidated, but variations in the degree of expression in the thymus have been implicated^{73,75}. Deficiencies in the complement component C1q are associated with SLE in a unique subset of families⁷⁶. The degree of penetrance of SLE in the absence of C1q is in excess of 90%, suggesting that a deficiency of this complement component is sufficient to mediate disease in a monogenic fashion. The precise molecular mechanism remains to be elucidated, although a role for C1q in early B cell development and the clearance of apoptotic cell bodies have been postulated^{77,78}. In this regard, several different complement deficiencies appear to potentiate SLE, strongly implicating this system in at least some pathogenic pathways to systemic autoimmunity^{79,80}.

A frame-shift variant and two missense mutations in *NOD2* have more recently been associated with susceptibility to Crohn's disease by TDT and case-control analysis^{81,82}. *NOD2* is located in the pericentromeric region of chromosome 16, a site previously shown in multiple linkage studies to contain a Crohn's susceptibility gene. The discovery of a chain-truncating frame-shift mutation in the *NOD2* gene, coupled with the knowledge that *NOD2* activates nuclear factor- κ B (NF- κ B) and confers responsiveness to bacterial lipopolysaccharides, a pathway long thought to be affected in Crohn's disease, strongly supports this inactive allele of *NOD2* as a

susceptibility allele in Crohn's disease. Finally, *AIRE* was recently identified by linkage and positional candidate analysis of affected Scandinavian populations^{60,61}.

Linkage studies in animal models have generally yielded statistical associations similar to those detected in human family studies; as a result, genomic locations have been equally imprecise. The most promising strategy for gene identification in animal models has been the classic strategy of congenic dissection, an approach pioneered by George Snell over 50 years ago⁸³. Congenic dissection separates the multiple genes mediating a polygenic autoimmune disease into a collection of congenic strains, each carrying one susceptibility gene on a single genetic background. Subsequent analysis of the component phenotypes expressed in each of the resultant congenic strains potentially allows a detailed characterization of the disease component contributed by each susceptibility gene in the original mouse strain and provides a phenotype to utilize for fine-mapping. Several investigators are following this strategy to characterize individual genes in animal models of IMD, SLE, MS and RA^{28,30,84-89}.

The identification of individual susceptibility genes is still ongoing, but several susceptibility alleles have been fine-mapped into extremely small congenic intervals, some of which contain strong positional candidates^{84,86,90}. The candidacy of *Il2* as *Idd3* is the most extensively investigated in positional candidate analysis to date^{91,92}. The *Il2* allele



in NOD mice differs from that in C57BL/10 mice by a complex mutation that results in a change in glycosylation that has been postulated to affect the *in vivo* function of interleukin 2 (IL-2)⁹². Although these structural changes in IL-2 are intriguing, conclusive evidence of a functional impact on the immune system by this *Ii2* polymorphism has not yet been reported. Congenic analysis has identified β_2 -microglobulin as a candidate for the *Idd13* locus on chromosome 2⁹⁰. A functional polymorphism between the alleles in NOR and NOD mice potentially affects MHC class I function and has been postulated to influence susceptibility to diabetes. Finally, *Sle1* on chromosome 1 and *Idd9* on chromosome 4 have each been dissected into multiple susceptibility loci *via* the creation of congenic recombinant chromosomes; several interesting positional candidates are in the process of being characterized^{86,93,94}.

Are susceptibility genes shared between AIDs?

Detailed analyses of linkage data in humans and rodent models have provided support for the hypothesis that common genes or genetic pathways may contribute to immune dysregulation and susceptibility to multiple AID. The basic data supporting this hypothesis are the observed colocalization of susceptibility loci in genome-wide scans in both mouse and human studies^{9,95-97}. Twelve separate non-MHC autoimmune susceptibility clusters have been identified in independent human and mouse genome scans (Fig. 2). These findings are intriguing and provide support for the idea that susceptibility to multiple AID may have some common susceptibility alleles or pathways. One caveat in interpretation is that the precision of this analysis is dependent upon the accuracy of allele placement by linkage studies, which is quite imprecise with multifactorial traits such as susceptibility to AID.

An alternative interpretation of the genetic colocalization of susceptibility alleles is that many immunologic genes are loosely clustered in mammalian genomes. In this interpretation, the colocalization of susceptibility alleles reflects disease associations with different genes in the same cluster, rather than a common allele. Data emerging from fine-mapping studies of susceptibility intervals from several animal models of AID support this alternative interpretation. The majority of the susceptibility segments that have undergone fine-mapping analysis in animal models have been found to contain more than one susceptibility locus^{86,90,94,98-102}. *Sle1*, for example, was recently split into four separate susceptibility loci, each contributing a unique aspect of the "SLE1" phenotype originally defined in congenic dissection. Fine-mapping studies of *Idd10*, *Idd13*, *Idd5*, *Idd1* and *Idd9* have resulted in the detection of a cluster of susceptibility loci within each of these congenic intervals.

The frequent detection of the genomic clustering of susceptibility genes may be interpreted in two ways. It is possible that the frequency of apparent clustering represents an ascertainment bias introduced by the relatively weak statistical power of the linkage analysis of multifactorial traits. That is, linkage is detectable only in regions of the genome that contain, by chance, several closely linked susceptibility alleles, the combined phenotypic impact of which yields a strong signal.

Alternatively, the detection of genomic clusters may represent an organizational feature of mammalian genomes in which genes involved in fundamental immune system pathways are occasionally clustered in specific regions, similar to the gene clusters associated with cytokine production⁹⁶. Data from the Human Genome Project has suggested the presence of functionally related gene clusters throughout the genome¹⁰³. The *Cia3* locus on murine chromosome 6

may provide an example of such clustering. This region is syntenic to regions on both human chromosome 12p12-p13 and rat chromosome 4 that contain susceptibility loci for IMD, SLE, MS and RA, and contains a cluster of attractive candidate genes, including *Tnfrsf1a*, *Ii5ra*, *Cd4*, *Cd27*, *Tgfa* and *Bphs*⁸⁷.

Thus, data supporting the organization of susceptibility alleles into genomic clusters continue to accumulate. These may, however, be interpreted as a mundane statistical anomaly or an intriguing new insight into the organization of mammalian genomes. Although a complete understanding must await a comparison of the disease alleles identified in multiple AIDs, it is reasonable to conclude that linkage colocalization reflects an important feature of the genetics of AID susceptibility. Certainly, regions of the genome that contain clusters of susceptibility genes warrant a top priority in future genomic analyses.

Future prospects

A crucial goal for future efforts in the genetics of AID will be the transition from linkage analysis and modeling into gene identification and disease pathway analysis. This process has been impeded by several factors, including the complexity of the mode of inheritance and the recently discovered genomic clustering of susceptibility genes. Success in the identification of susceptibility alleles in human populations will probably await the development of more powerful analytical procedures and larger patient populations. The extensive development of SNP technology may also facilitate gene identification, although SNPs alone may not suffice to overcome the complexities introduced into the analysis by epistasis and genetic heterogeneity.

The potential for success in AID gene identification in animal models is, on the other hand, excellent. Congenic dissection is a powerful tool that allows the characterization of phenotypes conferred by the individual genetic components of a polygenic disease, and standard fine-mapping procedures have been used to successfully narrow the susceptibility interval to as little as 800-1000 kb^{84,86}. The completion of the Human Genome Project and the soon-to-be-completed mouse project will provide quality molecular and physical maps for both of these mammalian genomes, which will greatly facilitate positional cloning efforts by providing rapid access and a precise localization of markers and positional candidate genes. In animal models, definitive gene involvement can be obtained by *in vivo* complementation using bacterial artificial chromosome (BAC) transgenic technologies^{104,105}. Targeted mutagenesis in BACs and/or embryonic stem cells can also help definitively to identify specific positional candidates as susceptibility genes.

Finally, the use of gene expression microarrays to identify genes whose expression is modified by AID or specific AID susceptibility alleles has the potential to revolutionize mapping strategies for complex traits. In theory, gene expression analysis can be used to delineate a plethora of component phenotypes in individuals with disease and their relatives, potentially providing a variety of new mapping strategies for complex traits. In addition, this technology should identify genes that are dysregulated because of the susceptibility allele, thus providing new insights into disease mechanisms and expanding the array of potential targets for the development of therapeutic strategies. This technology may also identify molecular expression phenotypes that may improve the identification of individuals who are at risk of developing disease, affording them the opportunity of preventive health care measures. Thus, although the analysis of multifactorial traits and AID in particular has been challenging, recent technical developments support an optimistic view of future developments.



1. Vyse, T.J. & Todd, J.A. Genetic analysis of autoimmune disease. *Cell* **85**, 311–318 (1996).
2. Rose, N.R. & Mackay, I.R. (eds) *The Autoimmune Diseases*. (Academic Press London, 1999).
3. Silman, A.J. et al. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. *Br. J. Rheumatol.* **32**, 903–907 (1993).
4. Winchester, R. in *Systemic lupus erythematosus* (ed. Lahita, R.G.) 65–85 (Churchill Livingstone, New York, 1992).
5. Davies, J.L. et al. A genome-wide search for human type 1 diabetes susceptibility genes. *Nature* **371**, 130–136 (1994).
6. Moser, K.L. et al. Genome scan of human systemic lupus erythematosus: evidence for linkage on chromosome 1q in african-american pedigrees. *Proc. Natl Acad. Sci. USA* **95**, 14869–14874 (1998).
7. Mein, C.A. et al. A search for type 1 diabetes susceptibility genes in families from the United Kingdom. *Nature Genet.* **19**, 297–300 (1998).
8. Rowe, R.E. et al. Linkage and association between insulin-dependent diabetes mellitus (IDDM) susceptibility and markers near the glucokinase gene on chromosome 7. *Nature Genet.* **10**, 240–242 (1995).
9. Jawaheer, D. et al. A genome-wide screen in multiplex rheumatoid arthritis families suggests genetic overlap with other autoimmune diseases. *Am. J. Hum. Genet.* **68**, 927–936 (2001).
10. Gaffney, P.M. et al. A genome-wide search for susceptibility genes in human systemic lupus erythematosus sib-pair families. *Proc. Natl Acad. Sci. USA* **95**, 14875–14879 (1998).
11. Shai, R. et al. Genome-wide screen for systemic lupus erythematosus susceptibility genes in multiplex families. *Hum. Mol. Genet.* **8**, 639–644 (1999).
12. Luo, D.F. et al. Confirmation of three susceptibility genes to insulin-dependent diabetes mellitus: IDDM4, IDDM5 and IDDM8. *Hum. Mol. Genet.* **5**, 693–698 (1996).
13. Concannon, P. et al. A second-generation screen of the human genome for susceptibility to insulin-dependent diabetes mellitus. *Nature Genet.* **19**, 292–296 (1998).
14. Sawcer, S. et al. A genome screen in multiple sclerosis reveals susceptibility loci on chromosome 6p21 and 17q22. *Nature Genet.* **13**, 464–468 (1996).
15. Ebers, G.C. et al. A full genome search in multiple sclerosis. *Nature Genet.* **13**, 472–476 (1996).
16. Salmon, J.E. et al. Fc γ RIIA alleles are heritable risk factors for lupus nephritis in African Americans. *J. Clin. Invest.* **97**, 1348–1354 (1996).
17. Leech, N.J. et al. Genetic and immunological markers of insulin dependent diabetes in Black Americans. *Autoimmunity* **22**, 27–32 (1995).
18. Lindqvist, A.K. et al. A susceptibility locus for human systemic lupus erythematosus (hSLE1) on chromosome 2q. *J. Autoimmunity* **14**, 169–178 (2000).
19. Rozzo, S.J., Vyse, T.J., Drake, C.G. & Kotzin, B.L. Effect of genetic background on the contribution of New Zealand black loci to autoimmune lupus nephritis. *Proc. Natl Acad. Sci. USA* **93**, 15164–15168 (1996).
20. Rozzo, S.J., Vyse, T.J., Menze, K., Izui, S. & Kotzin, B.L. Enhanced susceptibility to lupus contributed from the nonautoimmune C57BL/10, but not C57BL/6, genome. *J. Immunol.* **164**, 5515–5521 (2000).
21. Holmdahl, R. Genetics of susceptibility to chronic experimental encephalomyelitis and arthritis. *Curr. Opin. Immunol.* **10**, 710–717 (1998).
22. Jirholt, J. et al. Genetic linkage analysis of collagen-induced arthritis in the mouse. *Eur. J. Immunol.* **28**, 3321–3328 (1998).
23. Prochazka, M., Serreze, D.V., Frankel, W.N. & Leiter, E.H. NOR/Lt mice: MHC-matched diabetes-resistant control strain for NOD mice. *Diabetes* **41**, 98–106 (1992).
24. McDuffie, M. Derivation of diabetes-resistant congenic lines from the nonobese diabetic mouse. *Clin. Immunol.* **96**, 119–130 (2000).
25. McDuffie, M. Genetics of autoimmune diabetes in animal models. *Curr. Opin. Immunol.* **10**, 704–709 (1998).
26. Hogarth, M.B. et al. Multiple lupus susceptibility loci map to chromosome 1 in BXSB mice. *J. Immunol.* **161**, 2753–2761 (1998).
27. Gray, M. et al. Genome scan of human systemic lupus erythematosus by regression modeling: evidence of linkage and epistasis at 4p16–15.2. *Am. J. Hum. Genet.* **67**, 1460–1469 (2000).
28. Prins, J.-B. et al. Linkage on chromosome 3 of autoimmune diabetes and defective Fc receptor for IgG in NOD mice. *Science* **260**, 695–698 (1993).
29. Sundvall, M. et al. Identification of murine loci associated with susceptibility to chronic experimental autoimmune encephalomyelitis. *Nature Genet.* **10**, 313–317 (1995).
30. Morel, L. et al. Genetic reconstitution of systemic lupus erythematosus immunopathology with polygenic murine strains. *Proc. Natl Acad. Sci. USA* **97**, 6670–6675 (2000).
31. Morel, L., Tian, X.-H., Croker, B.P. & Wakeland, E.K. Epistatic modifiers of autoimmunity in a murine model of lupus nephritis. *Immunity* **11**, 131–139 (1999).
32. Morel, L., Rudofsky, U.H., Longmate, J.A., Schifflbauer, J. & Wakeland, E.K. Polygenic control of susceptibility to murine systemic lupus erythematosus. *Immunity* **1**, 219–229 (1994).
33. Kelley, V.E. & Winkelstein, A. Age- and sex-related glomerulonephritis in New Zealand white mice. *Clin. Immunol. Immunopathol.* **16**, 142–150 (1980).
34. Ghosh, S. et al. Polygenic control of autoimmune diabetes in nonobese diabetic mice. *Nature Genet.* **4**, 404–409 (1993).
35. Wu, J.M., Longmate, J.A., Adams, G., Hargrave, P.A. & Wakeland, E.K. Interval mapping of quantitative trait loci controlling humoral immunity to exogenous antigens: evidence that non-MHC immune response genes may also influence susceptibility to autoimmunity. *J. Immunol.* **157**, 2498–2505 (1996).
36. Bickerstaff, M.C. et al. Serum amyloid P component controls chromatin degradation and prevents antinuclear autoimmunity. *Nature Med.* **5**, 694–697 (1999).
37. Bolland, S. & Ravetch, J.V. Spontaneous autoimmune disease in Fc(γ)RIIB-deficient mice results from strain-specific epistasis. *Immunity* **13**, 277–285 (2000).
38. Botto, M. et al. Homozygous C1q deficiency causes glomerulonephritis associated with multiple apoptotic bodies. *Nature Genet.* **19**, 56–59 (1998).
39. Fossati, L. et al. An MRL/MpJ-lpr/lpr substrate with a limited expansion of lpr double-negative T cells and a reduced autoimmune syndrome. *Int. Immunol.* **5**, 525–532 (1993).
40. Warren, R.W., Caster, S.A., Roths, J.B., Murphy, E. & Pisetsky, D.S. The influence of the lpr gene on B cell activation: differential antibody expression in lpr congenic mouse strains. *Clin. Immunol. Immunopathol.* **31**, 65–77 (1984).
41. Kurtzke, J.F. Epidemiologic contributions to multiple sclerosis: an overview. *Neurology* **30**, 61–79 (1980).
42. Castaño, L. & Eisenbarth, G.S. Type-1 diabetes: a chronic autoimmune disease of human, mouse, and rat. *Annu. Rev. Immunol.* **8**, 647–679 (1990).
43. Kotzin, B.L. & O'Dell, J.R. in *Samter's Immunologic Diseases* (eds Frank, M.M., Austen, K.F., Claman, H.N. & Unanue, E.R.) 667–697 (Little, Brown, Boston, 1995).
44. Wicker, L.S., Todd, J.A. & Peterson, L.B. Genetic control of autoimmune diabetes in the NOD mouse. *Annu. Rev. Immunol.* **13**, 179–200 (1995).
45. Theofilopoulos, A.N., Koller, R., Singer, P. & Dixon, F.J. Molecular genetics of murine lupus models. *Adv. Immunol.* **46**, 61–109 (1989).
46. Kotzin, B. Systemic lupus erythematosus. *Cell* **85**, 303–306 (1996).
47. Willer, C.J. & Ebers, G.C. Susceptibility to multiple sclerosis: interplay between genes and environment. *Curr. Opin. Neurol.* **13**, 241–247 (2000).
48. Kurtzke, J.F. & Delasnerie, L. Reflection on the geographic distribution of multiple sclerosis in France. *Acta Neurol. Scand.* **93**, 110–117.
49. Kalman, B. & Lublin, F.D. The genetics of multiple sclerosis. A review. *Biomed. Pharmacother.* **53**, 358–370 (1999).
50. Harley, J.B. & James, J.A. Epstein-Barr virus infection may be an environmental risk factor for systemic lupus erythematosus in children and teenagers. *Arthritis Rheum.* **42**, 1782–1783 (1999).
51. James, J.A. et al. An increased prevalence of Epstein-Barr virus infection in young patients suggests a possible etiology for systemic lupus erythematosus. *J. Clin. Invest.* **100**, 3019–3026 (1997).
52. Todd, J.A. A protective role of the environment in the development of type 1 diabetes? *Diabetes Med.* **8**, 906–910 (1991).
53. Eisenberg, R.A., Craven, S.Y., Warren, R.W. & Cohen, P.L. Stochastic control of anti-Sm autoantibodies in MRL/MpJ-lpr/lpr mice. *J. Clin. Invest.* **80**, 691–697 (1987).
54. Gelehrter, T.D. & Collins, F.S. *Principles of Medical Genetics*. (Williams and Wilkins, Baltimore, MD, 1990).
55. Vwright, S. An analysis of variability in number of digits in an inbred strain of guinea pigs. *Genetics* **19**, 506–536 (1933).
56. Risch, N., Ghosh, S. & Todd, J.A. Statistical evaluation of multiple-locus linkage data in experimental species and its relevance to human studies: application to nonobese diabetic (NOD) mouse and human insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.* **53**, 702–714 (1993).
57. McAleer, M.A. et al. Crosses of NOD mice with the related NON strain: a polygenic model for type 1 diabetes. *Diabetes* **44**, 1186–1195 (1995).
58. Drake, C.G. et al. Analysis of the New Zealand Black contribution to lupus-like renal disease. Multiple genes that operate in a threshold manner. *J. Immunol.* **154**, 2441–2447 (1995).
59. Vyse, T.J., Todd, J.A. & Kotzin, B.L. in *The Autoimmune Diseases* (eds Rose, N.R. & Mackay, I.R.) 85–118 (Academic Press, London, 1998).
60. An autoimmune disease, APECED, caused by mutations in a novel gene featuring two PHD-type zinc-finger domains. The Finnish-German APECED Consortium. Autoimmune Polyendocrinopathy-Candidiasis-Ectodermal Dystrophy. *Nature Genet.* **17**, 399–403 (1997).
61. Nagamine, K. et al. Positional cloning of the APECED gene. *Nature Genet.* **17**, 393–398 (1997).
62. Eaves, I.A. et al. The genetically isolated populations of Finland and Sardinia may not be a panacea for linkage disequilibrium mapping of common disease genes. *Nature Genet.* **25**, 320–323 (2000).
63. Venter, J.C. et al. The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
64. Halushka, M.K. et al. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nature Genet.* **22**, 239–247 (1999).
65. Cargill, M. et al. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nature Genet.* **22**, 231–238 (1999).
66. Spielman, R.S. & Ewens, W.J. The TDT and other family-based tests for linkage disequilibrium and association. *Am. J. Hum. Genet.* **59**, 983–989 (1996).
67. Marron, M.P. et al. Genetic and physical mapping of a type 1 diabetes susceptibility gene (IDDM12) to a 100-kb phagemid artificial chromosome clone containing D2S72-CTLA4-D2S105 on chromosome 2q33. *Diabetes* **49**, 492–499 (2000).
68. Marron, M.P. et al. Insulin-dependent diabetes mellitus (IDDM) is associated with CTLA4 polymorphisms in multiple ethnic groups. *Hum. Mol. Genet.* **6**, 1275–1282 (1997).
69. Criswell, L.A. et al. PARP alleles and SLE: failure to confirm association with disease susceptibility. *J. Clin. Invest.* **105**, 1501–1502 (2000).
70. Bell, G.I., Horita, S. & Karam, J.H. A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. *Diabetes* **33**, 176–183 (1984).
71. Bell, G.I. et al. Recessive inheritance for the insulin linked IDDM predisposing gene. *Am. J. Hum. Genet.* **37**, 188 (1984).
72. Bennett, S.T. et al. Insulin VNTR allele-specific effect in type 1 diabetes depends on identity of untransmitted paternal allele. The IMDIAB Group. *Nature Genet.* **17**, 350–352 (1997).
73. Pugliese, A. et al. The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes. *Nature Genet.* **15**, 293–297 (1997).
74. Paquette, J., Giannoukakis, N., Polychronakos, C., Vafiadis, P. & Deal, C. The INS 5' variable number of tandem repeats is associated with IGF2 expression in humans. *J. Biol. Chem.* **273**, 14158–14164 (1998).
75. Vafiadis, P., Grabs, R., Goodyer, C.G., Colle, E. & Polychronakos, C. A functional analysis of the role of IGF2 in IDDM2-encoded susceptibility to type 1 diabetes. *Diabetes* **47**, 831–836 (1998).
76. Walport, M.J., Davies, K.A. & Botto, M. C1q and systemic lupus erythematosus. *Immunobiology* **199**, 265–285 (1998).
77. Carroll, M.C. The lupus paradox. *Nature Genet.* **19**, 3–4 (1998).
78. Taylor, P.R. et al. A hierarchical role for classical pathway complement proteins in the clearance of apoptotic cells in vivo. *J. Exp. Med.* **192**, 359–366 (2000).
79. Quigg, R. et al. Immune complex glomerulonephritis in C4- and C3-deficient mice. *Kidney Int.* **53**, 320–330 (1998).
80. Prodeus, A.P. et al. A critical role for complement in maintenance of self-tolerance. *Immunity* **9**, 721–731 (1998).
81. Ogura, Y. et al. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* **411**, 603–606 (2001).
82. Hugot, J.P. et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* **411**, 599–603 (2001).
83. Snell, G.D. Methods for the study of histocompatibility genes. *J. Genetics* **49**, 87 (1948).
84. Lyons, P.A. et al. Congenic mapping of the type 1 diabetes locus, *Idd3*, to a 780-kb region of mouse chromosome 3: identification of a candidate segment of ancestral DNA by haplotype mapping. *Genome Res.* **10**, 446–453 (2000).
85. Morel, L., Mohan, C., Croker, B.P., Tian, X.-H. & Wakeland, E.K. Functional dissection of systemic lupus erythematosus using congenic mouse strains. *J. Immunol.* **158**, 6019–6028 (1997).
86. Morel, L., Blenman, K.R., Croker, B.P. & Wakeland, E.K. The major murine systemic lupus erythematosus susceptibility locus, *Sle1*, is a cluster of functionally related genes. *Proc. Natl Acad. Sci. USA* **98**, 1787–1792 (2001).
87. McIntoe, R.A. et al. Localization of non-Mhc collagen-induced arthritis susceptibility loci in DBA/1J mice. *Proc. Natl Acad. Sci. USA* **96**, 2210–2214 (1999).
88. Lundberg, C., Lidman, O., Holmdahl, R., Olsson, T. & Piehl, F. Neurodegeneration and glial activation patterns after mechanical nerve injury are differentially regulated by non-MHC genes in congenic inbred rat strains. *J. Comp. Neurol.* **431**, 75–87 (2001).
89. Morel, L., Yu, Y., Blenman, K.R., Caldwell, R.A. & Wakeland, E.K. Production of congenic mouse strains carrying SLE-susceptibility genes derived from the SLE-prone NZM/Aeg2410 strain. *Mamm. Genome* **7**, 335–339 (1996).
90. Serreze, D.V. et al. Subcongenic analysis of the *Idd13* locus in NOD/Lt mice: evidence for several susceptibility genes including a possible diabetogenic role for β 2-microglobulin. *J. Immunol.* **160**, 1472–1478 (1998).
91. Lord, C.J. et al. Mapping the diabetes polygene *Idd3* on mouse chromosome 3 by use of novel congenic strains. *Mamm. Genome* **6**, 563–570 (1995).



92. Podolin, P. L. *et al.* Differential glycosylation of interleukin 2, the molecular basis for the NOD *Idd3* type 1 diabetes gene? *Cytokine* **12**, 477–482 (2000).
93. Siegmund, T. *et al.* Analysis of the mouse CD30 gene: a candidate for the NOD mouse type 1 diabetes locus *Idd9.2*. *Diabetes* **49**, 1612–1616 (2000).
94. Lyons, P. A. *et al.* The NOD *Idd9* genetic interval influences the pathogenicity of insulinitis and contains molecular variants of Cd30, Tnfr2, and Cd137. *Immunity* **13**, 107–115 (2000).
95. Becker, K. G. *et al.* Clustering of non-MHC susceptibility candidate loci in human autoimmune diseases. *Proc. Natl Acad. Sci. USA* **95**, 9979–9984 (1998).
96. Becker, K. G. Comparative genetics of type 1 diabetes and autoimmune disease: common loci, common pathways? *Diabetes* **48**, 1353–1358 (1999).
97. Griffiths, M. M., Encinas, J. A., Remmers, E. F., Kuchroo, V. K. & Wilder, R. L. Mapping autoimmunity genes. *Curr. Opin. Immunol.* **11**, 689–700 (1999).
98. Podolin, P. L. *et al.* Congenic mapping of the insulin-dependent diabetes (*Idd*) gene, *Idd10*, localizes two genes mediating the *Idd10* effect and eliminates the candidate *Fcrg1*. *J. Immunol.* **159**, 1835–1843 (1997).
99. Cordell, H. J., Todd, J. A. & Lathrop, G. M. Mapping multiple linked quantitative trait loci in non-obese diabetic mice using a stepwise regression strategy. *Genet. Res.* **71**, 51–64 (1998).
100. Hill, N. J. *et al.* NOD *Idd5* locus controls insulinitis and diabetes and overlaps the orthologous CTLA4/IDDM12 and NRAMP1 loci in humans. *Diabetes* **49**, 1744–1747 (2000).
101. Hattori, M. *et al.* Cutting edge: homologous recombination of the MHC class I K region defines new MHC-linked diabetogenic susceptibility gene(s) in nonobese diabetic mice. *J. Immunol.* **163**, 1721–1724 (1999).
102. Podolin, P. L. *et al.* Localization of two insulin-dependent diabetes (*Idd*) genes to *Idd10* region on mouse chromosome 3. *Mamm. Genome* **9**, 283–286 (1998).
103. Caron, H. *et al.* The human transcriptome map: clustering of highly expressed genes in chromosomal domains. *Science* **291**, 1289–1292 (2001).
104. Antoch, M. P. *et al.* Functional identification of the mouse circadian *clock* gene by transgenic BAC rescue. *Cell* **89**, 655–667 (1997).
105. Probst, F. J. *et al.* Correction of deafness in shaker-2 mice by an unconventional myosin in a BAC transgene. *Science* **280**, 1447–1451 (1998).
106. Redondo, M. J. *et al.* Genetic determination of islet cell autoimmunity in monozygotic twin, dizygotic twin, and non-twin siblings of patients with type 1 diabetes: prospective twin study. *Br. Med. J.* **318**, 698–702 (1999).
107. Chan, O. T., Madaio, M. P. & Shlomchik, M. J. B cells are required for lupus nephritis in the polygenic, Fas-intact MRL model of systemic autoimmunity. *J. Immunol.* **163**, 3592–3596 (1999).
108. Gulko, P. & Winchester, R. *Lupus: Molecular and cellular pathogenesis* (eds Kammer, G. & Tsokos, G. C.) 101–123 (Humana Press, Totowa, 1999).
109. She, J. X. & Marron, M. P. Genetic susceptibility factors in type 1 diabetes: linkage, disequilibrium and functional analyses. *Curr. Opin. Immunol.* **10**, 682–689 (1998).