Genetic background for immune-mediated diseases

Frode Vartdal

Institute of Immunology, Rikshospitalet, The National Hospital, University of Oslo, Oslo, Norway

Vartdal F. Genetic background for immune-mediated diseases. Acta Odontol Scand 2001;59:212–215. Oslo. ISSN 0001-6357.

Gene variants (alleles) involved in the immune response are most likely selected during evolution. The allelic polymorphisms that may be advantageous in fighting harmful agents may be susceptibility genes in immune-mediated diseases. Identification of susceptibility genes is important because these genes encode proteins, which are most probably involved in the disease process. Hence, the identification of susceptibility genes may lead to an improved understanding of the pathogenesis and may therefore help the development of preventive and therapeutic measures. Susceptibility genes may be identified by analyzing genes known to be involved in immune responses (candidate gene search) or by analyzing gene markers evenly distributed over the genome (genome-wide scan). However, since several genes jointly contribute to disease susceptibility, the frequencies of single susceptibility genes may be quite high in the normal population. Moreover, different set of genes may predispose to the same clinical disease. It may therefore be very difficult to identify susceptibility genes, apart from the major histocompatibility complex (MHC) genes, which have now been shown to predispose to several immune-mediated diseases. □ *Disease susceptibility; genes; major histocompatibility complex; markers; polymorphisms*

Frode Vartdal, Institute of Immunology, Rikshospitalet, The National Hospital, NO-0027 Oslo, Norway. Tel: +47 23071343, fax: +47 23073510.

The immune system protects us against pathogenic infectious agents. During a protective response the immune system may induce disease symptoms by destroying cells locally or by producing soluble mediators (cytokines) with systemic effects. In some instances the immune system may be the main or sole factor contributing to disease. Such immune-mediated diseases include disease entities such as allergies, autoimmune diseases, celiac disease, and periodontitis. In many of these diseases the immunopathogenesis is poorly understood. Since disease susceptibility genes encode proteins, which are most probably involved in the disease process, identification of susceptibility genes may point directly to proteins participating in the disease process and thus help to unveil its molecular mechanisms (1). This may be important for the development prophylactic and therapeutic measures.

The immune system must be able to detect and destroy a wide variety of infectious agents. During evolution the host species has generated variants (alleles) of genes involved in the immune response, in order for the species to survive the enormous pressure applied by pathogenic infectious organisms. Since each species is endowed with several allelic genetic variants, there will always be some individuals in a species who will survive any plague. Unfortunately, allelic variants, which may be superior immune-response genes during an infection, could turn out to predispose for immune-mediated diseases. Thus, there is reason to believe that the genes that predispose for such diseases are normal variants of genes selected for during evolution (2, 3). This paper will briefly discuss how and why immune-genes may predispose for immune-mediated diseases and, finally, present some strategies and techniques for the identification of such genes.

The importance of genetic and environmental factors

Familial aggregation of a disease suggests that genes and/or environmental factors contribute to the disease (4). To assess the genetic and environmental components, the concordance rates of monozygotic (MZ) twins are compared with those of dizygotic (DZ) twins. In autoimmune diseases the concordance rates of MZ twins are between 20% and 50%, whereas the concordance rates of DZ twins are usually much lower (typically 1%–5%)(5–7). Since both MZ and DZ twins share the same milieu, these data indicate that genetic factors are important for the development of such diseases. Since MZ twins share the same genes, the susceptibility not conferred by genes has been attributed to environmental factors. However, in assessing the importance of environmental factors, it is important to be aware that MZ twins cease to be genetically identical as soon as the immune system starts to develop. Thus, during the generation of the antigen-specific receptors on T and B lymphocytes the genetic elements encoding the variable part of the T-cell receptors (TCR) and B-cell receptors (BCR) undergo somatic recombination and somatic mutations (only BCR genes) (8, 9). This is a stochastic process that in genetically predisposed individuals may generate high-affinity T and B lymphocytes,

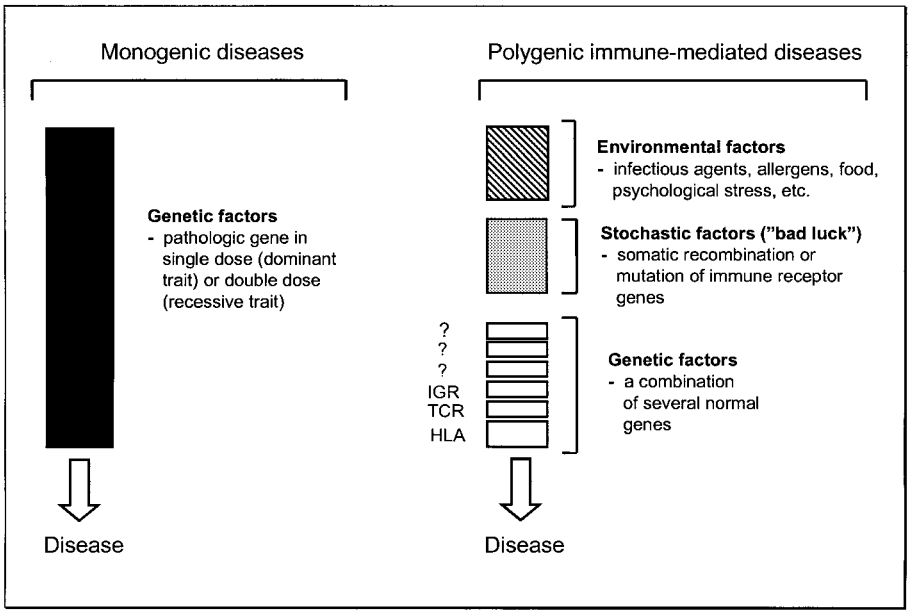

Fig. 1. Factors determining the disease susceptibility in monogenic diseases and in polygenic diseases like the immune-mediated diseases discussed in this paper.

which on recognition of self-antigens or external antigens could cause immune-mediated diseases (10).

Multiple normal genes confer susceptibility to immune-mediated diseases

In immune-mediated diseases the observation that there is a sharp fall in recurrence risk from MZ to other first-degree relatives suggests that there are several susceptibility genes that predispose to the same disease—that is, they are poly-or multi-genic diseases (11). In monogenic diseases, on the other hand, one gene in single (dominant) or double (recessive) dose determines the disease susceptibility (Fig. 1). Susceptibility genes in monogenic diseases are abnormal genes, which either encode a dysfunctional protein (hypo-or hyper-functional) or fail to encode a protein. For a long time immunologists searched for disease-unique dysfunctional susceptibility genes also in immune-mediated diseases. However, the susceptibility genes identified so far in immune-mediated diseases appear to be completely normal genes (2, 3). Thus, it seems that the genetic susceptibility for these diseases is conferred by an unfortunate combination of normal genes.

Different sets of genes may predispose to the same disease

Immune-mediated diseases may be heterogenous on the molecular level, although they display similar clinical and pathologic anatomical features. Thus, different major histocompatibility complex (MHC) genes/molecules may predispose to the same disease (12, 13). The reason is probably that the antigen-binding pockets of MHC molecules can bind different peptides from the disease-inducing antigen (14) and present it to T cells (1). This is clearly seen in the animal models of multiple sclerosis, experimental allergic encephalomyelitis (EAE), in which animals with different MHC molecules bind different peptides from the encephalitogenic protein, myelin basic protein, and present it to different sets of T cells (15–17). It is also possible that certain susceptibility genes may be dispensable if an individual carries high numbers of other susceptibility genes and is exposed to a high load of environmental factors.

Strategies for the identification of susceptibility genes

Identification of susceptibility genes in polygenic diseases is an arduous task. First, the fact that some susceptibility genes may not always be necessary for the development of disease may profoundly disturb the genetic analysis of patient materials. Second, high frequencies of susceptibility alleles in the normal population that is examined make it difficult to show a significant association or linkage of a disease to true susceptibility genes. Third, confounding factors in a heterogeneous patient or healthy control material (such as recent ethnic admixture) may lead to the observation of spurious associations.

In monogenic diseases susceptibility genes are usually identified by the help of *linkage studies*, by which a genetic linkage is shown to exist when a genetic marker and the disease phenotype co-segregate in a family (18). However,

such conventional linkage studies will usually fail to identify linkage in polygenic immune-mediated diseases, in which susceptibility genes may not be necessary or may occur with too high frequency in the normal population to enable statistically valid evaluation in the relatively small family materials usually available. An alternative to classical linkage studies is to examine whether affected sib pairs more often than expected share alleles of a marker locus (25%, 50%, and 25% probability for sharing two, one, or zero alleles, respectively) (19). If the affected sibs share two and one alleles more often than expected, this suggests that the marker gene or a closely linked gene contribute to the genetic susceptibility. Such *identity by descent* (IBD) analysis is more robust than classical linkage studies, since only affected members of a family are included in the study.

When genes only confer susceptibility and are neither sufficient nor necessary for the development of disease, identification by *association studies* is as useful as or better than linkage studies in families (20). Association studies examine whether patients carry a given allele at a locus more frequently than healthy individuals of the same ethnic group. Identification of HLA (human MHC) susceptibility genes has mostly been performed by association studies (reviewed in Refs. 21, 22). This type of studies may display false-positive associations if the patients and the healthy controls are derived from a population with a recent admixture of other ethnic groups. To circumvent the problems of heterogeneity in and between the patient and control groups, family-based controls are often used. The most widely used test for association in families is the *transmission disequilibrium test* (TDT), which examines whether the transmission of alleles from heterozygous parents to affected children deviates from the expected 50% (23).

Susceptibilty genes in immune-mediated diseases may be identified by studying polymorphic genes known to be important in the immune response. This *candidate gene approach* has so far led to the identification of most of the susceptibility genes found in immune-mediated diseases (3, 12, 21, 22, 24, 25). Alternatively, tentative candidate genes may first be identified in experimental animal models of human disease, after which the identified susceptibility genes can be studied in humans (26).

With the advent of a comprehensive map of a large number of polymorphic gene markers and with new efficient techniques to screen for these polymorphisms, it has become feasible to perform so-called *genome-wide search* for markers of susceptibility genes (27–30). If a disease can be consistently shown to be linked to, or associated with, a genetic marker locus, one can use this as a starting point to identify the true susceptibility locus. This is possible because alleles at neighboring loci display linkage disequilibrium, which is the phenomenon in which the alleles of two loci occur more often together than expected if they segregated independently. Linkage disequilibrium is a prerequisite for successful positional cloning. However, it may often pose a serious problem for the identification of the primary susceptibility loci in gene complexes (such as HLA), in which a large set of immune genes is located. Studies of the susceptibility genes in several ethnic groups, which display variable linkage disequilibria, may then be helpful. Thus, maintenance of an association of a candidate gene across several ethnic groups with different linkage disequilibria in the genetic region studied strongly indicates that this gene is a true susceptibility gene (31, 32).

Polymorphic markers and techniques employed for the study of susceptibility genes

The human genome contains several polymorphic markers that can be used in the search for susceptibility genes. Such markers include the *single nucleotide polymorphisms* (SNP) (33) and repeat-sequence DNA, so-called '*variable number of tandem repeats*' (VNTR) (34). SNPs are the major source of polymorphisms in the genome and may be identified by a wide variety of techniques. If they are part of a target site for restriction enzymes, they may be identified by cleavage of the site (restriction fragment length polymorphism (RFLP) and polymerase chain reaction RFLP (PCR-RFLP) techniques). SNP can also be identified by probing with sequence-specific oligonuleotides (SSO), by sequence-specific PCR (SSP) amplification, or by sequencing. VNTRs are most easily identified by measuring their length with DNA sequencer machines (35). Studying several thousand VNTRs in large patient and control materials is a very work- and time-demanding venture. This work can now be greatly facilitated by VNTR analysis of pooled DNA from patients and comparison with pooled DNA from ethnically matched healthy controls (36).

After the publication of the entire humane genome, there will be a strong demand for efficient techniques for typing of genetic polymorphisms. The next few years will therefore without doubt see many new typing devices, which will greatly enhance capacity. This will provide us with enormous amounts of data. Since the number of markers analyzed (several thousands) is rapidly increasing, it may become very difficult to sort out true susceptibility polymorphisms from those displaying spurious associations. Therefore very large numbers of patients and controls will have to be studied. Moreover, linkages and associations observed in one study should always be confirmed in studies performed on independent samples, preferentially collected from another ethnic group than in the initial study. Collection and analysis of sufficiently large data sets from various ethnic groups may probably best be performed in joint international ventures, such as that performed by the Genetic Analysis of Multiple Sclerosis in EuropeanS (GAMES) study group (www.mrc-bsu.cam.ac.uk/MSgenetics/GAMES/).

Conclusions

Genetic susceptibility to immune-mediated diseases seems

to be conferred by an unfortunate combination of normal genes. Since each of these genes may be frequent in the normal population and not always necessary for the development of disease, they may be very difficult to trace. However, with the advent of comprehensive genetic maps and automatic gene typing techniques, it is believed that it will soon be possible to mine deeper into the field of immunogenetics in an attempt to unveil the pathogenesis of immune-mediated diseases.

References

1. Sollid LM. The molecular basis of celiac disease. Annu Rev Immunol 2000;18:53–81.
2. Todd JA, Bell JI, McDevitt HO. HLA-DQβ gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature 1987;329:599–604.
3. Dai KZ, Harbo HF, Celius EG, Otuari A, Sørensen PS, Ryder LP, et al. The T cell regulator gene *SH2D2A* contributes to the genetic susceptibility of multiple sclerosis. Genes Immunity. In press 2001.
4. Risch N. Assessing the role of HLA-linked and unlinked determinants of disease. Am J Hum Genet 1987;40:1–14.
5. Kaprio J, Tuomilehto J, Koskenkvuo M, Romanov K, Reunanen A, Eriksson J, et al. Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes in a population-based cohort of twins in Finland. Diabetologica 1992;35:1060–7.
6. Kyvik KO, Green A, Beck-Nielsen H. Concordance rates of insulin dependent diabetes mellitus: a population based study of young Danish twins. BMJ 1995;311:913–7.
7. Ebers GC, Bulman DE, Sadovnick AD, Paty DW, Warren S, Hader W, et al. A population-based study of multiple sclerosis in twins. N Engl J Med 1986;315:1638–44.
8. Tonegawa S. Somatic generation of antibody diversity. Nature 1983;302:575–81.
9. Davis MM, Bjorkman PJ. T cell antigen receptor genes and T cell recognition. Nature 1988;334:395–402.
10. Möller E, Böhme J, Valugerdi MA, Ridderstad A, Olerup O. Speculations on mechanisms of HLA associations with autoimmune diseases and the specificity of 'autoreactive' T lymphocytes. Immunol Rev 1990;118:5–19.
11. Risch N. Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 1990;46:222–8.
12. Sollid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E. Evidence for a primary association of celiac disease to a particular HLA-DQαβ heterodimer. J Exp Med 1989;169:345–50.
13. Spurkland A, Sollid LM, Polanco I, Vartdal F, Thorsby E. HLA-DR and –DQ genotypes of celiac disease patients serologically typed to be non-DR3 or non-DR5/7. Hum Immunol 1992;35:188–92.
14. Rammensee HG, Friede T, Stevanovic S. MHC ligands and peptide motifs: first listing. Immunogenetics 1995;41:178–228.
15. Zamvil SS, Mitchell DJ, Lee NE, Moore AC, Waldor MK, Sakai K, et al. T cell epitope of the autoantigen myelin basic protein that induces encephalomyelitis. Nature 1986;324:258–60.
16. Kono DH, Urban JL, Horvath SJ, Ando DG, Saavedra RA, Hood L. Two minor determinants of myelin basic protein induce experimental allergic encephalomyelitis in SJL/J mice. J Exp Med 1988;168:213–27.
17. Zamvil SS, Mitchell DJ, Powell MB, Sakai K, Rothbard JB, Steinman L: Multiple discrete encephalitogenic epitopes of the autoantigen myelin basic protein that include a determinant for I-E class II-restricted T-cells. J Exp Med 1988;168:1181–6.
18. The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell 1993;72:971–83.
19. Seboun E, Robinson MA, Doolitle TH, Ciulla TA, Kindt TJ, Hauser SL. A susceptibility locus for multiple sclerosis is linked to the T cell receptor β chain complex. Cell 1989;57:1095–100.
20. Greenberg DA. Linkage analysis of 'necesssary' disease loci versus 'susceptibility' loci. Am J Hum Genet 1993;52:135–43.
21. Tiwaris JL, Terasaki PI. HLA and disease associations. New York: Springer-Verlag; 1985.
22. Moffatt ME, Cookson WOCM. Allergy. In: Hall IP, editor. Genetics of asthma and atopy. Basel: Karger; 1996. p. 71–97.
23. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium; the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 1993;52:506–16.
24. Harbo HF, Celius EG, Vartdal F, Spurkland A. CTLA4 promoter and exon1 dimorphisms in multiple sclerosis. Tissue Antigens 1999;53:106–10.
25. Hill MR, Cookson WOCM. A new variant of the (subunit of the high-affinity receptor for immunoglobulin E (FcεRI-β E237G): associations with measures of atopy and bronchial hyperresponsiveness. Hum Mol Genet 1996;5:959–62.
26. Vyse TJ, Todd JA. Genetic analysis of autoimmune diseases. Cell 1996;85:311–8.
27. Davies JL, Kawaguchi Y, Bennett ST, Copeman JB, Cordell HJ, Pritchard LE, et al. A genome-wide search for human type 1 diabetes susceptibility genes. Nature 1994;371:130–6.
28. Ebers GC, Kukay K, Bulman DE, Sadovnick AD, Rice G, Anderson C. A full genome search in multiple sclerosis. Nat Genet 1996;13:472–6.
29. Wjst M, Fischer G, Immervoll T, Jung M, Saar K, Rueschendorf F, et al. A genome-wide search for linkage to asthma. German Asthma Genetics Group. Genomics 1999;58:1–8.
30. Satsangi J, Parkes M, Louis E, Hashimoto L, Kato N, Welsh K. Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. Nat Genet 1996;14:199–202.
31. Thorsby E, Ronningen KS. Particular HLA-DQ molecules play a dominant role in determining susceptibility or resistance to type 1 (insulin-dependent) diabetes mellitus. Diabetologia 1993;36:371–7.
32. Mignot E, Lin L, Rogers W, Honda Y, Qiu X, Lin X. Complex HLA-DR and -DQ interactions confer risk of narcolepsy-cataplexy in three ethnic groups. Am J Hum Genet 2001;68:686–99.
33. The International SNP Map Working Group. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 2001;409:928–33.
34. Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, et al. The 1993–94 Genethon human genetic linkage map. Nat Genet 1994;2:246–339.
35. Lie BA, Todd JA, Pociot F, Nerup J, Akselsen HE, Joner G. The predisposition to type 1 diabetes linked to the human leukocyte antigen complex includes at least one non-class II gene. Am J Hum Genet 1999;64:793–800.
36. Barcellos LF, Klitz W, Field LL, Tobias R, Bowcock AM, Wilson R, et al. Association mapping of disease loci, by use of a pooled DNA genomic screen. Am J Hum Genet 1997;61:734–47.