Incomplete penetrance of MHC susceptibility genes: prospective analysis of polygenic MHC-determined traits

Key words: major histocompatibility complex; incomplete penetrance; autoimmunity; polygenic disease

Acknowledgments: This work was supported by grants HL-29583 and AI-14157 from the National Institutes of Health. The authors especially thank Dr. Charles E. Larsen for his thoughtful comments. They also thank Drs. Devendra P. Dubey, Keith Crawford, Gabriel Uko, John Xu and Fred S. Rosen for their valuable suggestions.

Abstract: We propose an approach to understanding incomplete penetrance of disease susceptibility genes as a method of studying the underlying mechanisms of polygenic diseases. Incomplete penetrance is the failure of genetically susceptible individuals to exhibit a trait. We define as baseline penetrance that which occurs in genetically identical (monozygotic) twins of an index subject with a major histocompatibility complex (MHC)-associated disease or trait. We consider two mechanisms for incomplete baseline penetrance: an extrinsic (environmental) trigger and an intrinsic stochastic, gene-associated process. The latter can be detected for dominant expression because susceptibility genes in homozygotes (with their two intrinsic triggers) will be up to twice as frequently penetrant as those in heterozygotes. The extent of MHC and non-MHC gene contribution determines differences between baseline penetrance and apparent penetrance in MHC-identical sib pairs, sib pairs in general and MHC-identical unrelated individuals. Inheritance patterns in families do not reveal modes of inheritance of incompletely penetrant polygenic MHC-determined traits. A method is proposed to study such traits prospectively in persons presumed to be homozygous, heterozygous or non-carrying for susceptibility genes by determining trait expression in homozygotes, heterozygotes, and non-carriers of trait-associated conserved extended MHC haplotypes. The method provides direct estimates of apparent penetration rates, modes of genetic determination, and, if the trait is dominant, the origin of penetrance. When applied to dominant MHC susceptibility gene-determined immunoglobulin deficiencies in two populations, the ratios of affected haplotype homozygotes to heterozygotes near 2.0 were consistent with an intrinsic mechanism for baseline penetrance acting on the MHC susceptibility genes.

The burgeoning of molecular biological techniques over the past several decades has produced incredible advances in our understanding of human genetics. Despite this, progress in unraveling many complex (polygenic) diseases, particularly those with major histocompatibility complex (MHC) involvement, has been slow. There have been enormous efforts to identify and to understand the role of MHC and non-MHC susceptibility genes in autoimmune diseases.
diseases such as type 1 diabetes. Nevertheless, for most, we have yet to even identify unequivocally the responsible genes, let alone understand how they cause disease. At least part of the problem is incomplete penetrance of susceptibility genes.

Incomplete penetrance in its historic and broadest sense is the failure of some individuals who carry a susceptibility gene marker to express the trait (disease, phenotype, function) in question. For example, among Caucasian patients with type 1 diabetes, a classic HLA-associated disease, there is an increase in the frequency of HLA-DR3 and DR4 and particularly of HLA-DR3/DR4 heterozygotes. Nevertheless, most Caucasians with these HLA types do not have diabetes.

Incomplete penetrance is one of the elements of the “geneticist’s nightmare” (1) that characterizes many MHC-determined autoimmune diseases and phenomena. It confounds any direct segregation or linkage analysis because over 90% of affected persons have no affected family members (2).

We believe that new methods of theoretic genetic analysis to complement the rapid accumulation of molecular findings are necessary for understanding the mechanisms of polygenic diseases. In this paper, we shall analyze the elements of penetrance in MHC-associated disorders, separating the contributions of non-MHC susceptibility genes from the baseline penetrance observed in monozygotic twin pairs. We will then explore possible mechanisms of baseline penetrance, including an external environmental trigger and an intrinsic random process affecting the susceptibility gene rather than the whole subject. From these considerations, we will develop a prospective approach to determine the mode of inheritance and apparent penetrance in unrelated subjects homozygous, heterozygous and non-carrying for conserved extended MHC haplotypes known or suspected to carry susceptibility genes. Finally, data for immunoglobulin deficiencies studied in this way will be analyzed to show that, for dominant traits, the ratio of affected homozygotes to heterozygotes supports an intrinsic, MHC gene-based mechanism for baseline penetrance.

**Baseline and apparent penetrance**

It is important to distinguish two aspects of penetrance. The first is penetrance that is “baseline” as manifested by trait (disease) concordance rate in monozygotic (genetically identical) twins. The other is “apparent” penetrance that is observed in some defined group of subjects such as persons in the general population with genetic markers for the trait or defined first-degree relatives of a person with the trait. This distinction is easily illustrated by MHC-marker family studies of type 1 diabetes (2). In that disorder, baseline penetrance, as measured by disease concordance in monozygotic twins, is 30 to 50% (3). This rate would presumably apply to all fully susceptible persons in the population in general. Apparent penetrance in MHC-identical sibs of patients is on the order of 15%, in sibs in general is 5 to 6%, in parents or children is 3%, and in persons homozygous or doubly heterozygous for the diabetes-associated genetic markers HLA-DR3 and/or DR4 is about 3 to 5% (2).

These descending rates of apparent penetrance are characteristic of polygenic disease and reflect the descending frequencies of MHC and non-MHC susceptibility genes in these groups. A monozygotic twin of a patient has the same complete set of susceptibility genes as the patient, no matter how many susceptibility loci there are. An MHC-identical sib of a patient has the necessary MHC susceptibility genes but may or may not have the necessary non-MHC susceptibility genes. All family members have a higher frequency of MHC and non-MHC susceptibility genes than the general population because the family was ascertained through a patient who has all of them. A patient’s sibs in general have a lower rate than MHC-identical sibs because MHC susceptibility is not assured. Similarly, in a child or parent of a patient, one allele at each susceptibility locus has a frequency equivalent to the background population frequency. In summary, baseline penetrance is the frequency with which a monozygotic twin of an affected individual has the trait. Apparent penetrance varies according to the group of individuals specified and reflects both baseline penetrance and the frequencies of all susceptibility genes in that group.

**Penetrance and concordance**

Although it is usual to equate penetrance with concordance rates for an MHC-determined trait in monozygotic twin pairs, this is true only if one identifies the pairs through one affected member (4). In essence, one uses the affected twin to determine that the other twin is genetically completely susceptible. One then determines the prevalence of trait expression in the second twins as a group. This is the same as the rate of concordance for trait expression in the pairs ascertained through an affected twin. One can extend this approach to MHC-susceptible persons. For example, if one wishes to know the concordance rate in sibs MHC-identical (by descent) to the patient, one uses the patient to define MHC susceptibility. Apparent penetrance in this sib group is the same as concordance for pairs of patients and their MHC-identical sibs. In similar fashion, one can use the fixity of DNA on MHC susceptibility gene-bearing conserved extended haplotypes among random unrelated people to study trait expression, as we shall see below.
Mode of genetic determination

To explore methods of study of incompletely penetrant polygenic MHC-determined phenomena, we need first to consider some basic genetic aspects of such situations. For traits determined recessively by an MHC susceptibility gene, only homozygotes for these genes or closely linked marker genes will express the trait. If there is one such MHC allele (or marker), only homozygotes for the allele or marker are susceptible. If there is more than one, homozygotes for each such allele or marker as well as heterozygotes for any pair of susceptibility alleles or markers will be MHC-susceptible. To be fully susceptible for the trait in question, all other (non-MHC) susceptibility gene requirements must also be met in an aggregate susceptibility genotype with at least one susceptibility allele at each locus for a dominant trait and two for a recessive trait. Population penetrance is determined by the ratio of affected individuals (in general or with a specified MHC genotype) to all persons or to those with this genotype in a population. For a trait determined by a dominantly expressed MHC susceptibility gene, both homozygotes and heterozygotes for MHC susceptibility genes are susceptible but, again, non-MHC susceptibility genes must also be present in correct dose for the individual to be genetically susceptible.

The inadequacy of family studies

The high population frequency of susceptibility genes in polygenic MHC-associated phenomena and diseases is at least as important as incomplete baseline penetrance in obscuring the genetic expression of MHC-determined traits in family studies. For fully penetrant susceptibility genes that alone determine the trait, the patterns of inheritance within a family are distinctive for rare dominant and rare recessive genes. The rare dominant gene D causes disease in all three generations with half of people at risk for inheriting it (i.e., all those who carry it) expressing the trait. In contrast, a rare recessive susceptibility gene in a monogenic disease is inherited by and expressed in 25% of sibs of a patient.

These patterns break down when susceptibility genes or their markers are common. Fig. 1A shows how the pattern produced by a dominant gene is completely mimicked by the presence of multiple independent examples of fully penetrant recessive d genes. For incompletely penetrant susceptibility genes, this problem is exacerbated, as is seen in Fig. 1B and 1C, where affected persons happen to be present in two generations, as well as among sibs, for common or rare dominant and common recessive susceptibility genes at approximately 50% penetrance.

It is clear that methods of analysis other than examination of inheritance patterns in families must be used to define modes of genetic determination. One might also hope that some different method of analysis than is currently available might provide insight into the mechanism of baseline penetrance.

Retrospective approaches

Most MHC allele-association studies have been retrospective in the sense that patients ascertained for a disease or trait are treated as a population and analyzed for the frequency-distributions of MHC alleles in comparison with an ethnically matched healthy control population. From family studies of patients, haplotypes, including conserved extended haplotypes and their fragments, can be recognized with certainty (5–10). Moreover, a set of control MHC haplotypes, those not occurring in any patient in the family, which we call family control haplotypes (11), can be obtained. One can test statistically for any differences in frequency in patient and control haplotypes or populations and one can calculate an odds ratio for any marker or whole haplotype. Except as noted below, such studies provide information about neither mode of inheritance nor penetrance.

There are two main currently available retrospective approaches to decide the mode of genetic determination of traits caused by partially penetrant MHC susceptibility genes. One can analyze affected sib pairs (12). If the trait and its genetic markers are rare, in a recessive disorder all or most sib pairs will be identical for the marker or markers on both chromosomes. In a rare dominant disorder in which one parent carries a susceptibility gene, they need share one marker only and the sibs will be MHC haplidential or MHC identical in equal proportions. Since many of the traits (diseases) of interest are relatively common and susceptibility genes are very common, however, such clear-cut results are usually not obtained. For example, in type 1 diabetes, about 55% of affected sib pairs are MHC identical, about 38% are haplidential and about 7% have no MHC haplotypes in common (13). This distribution does not fit dominant inheritance but is actually consistent with recessive inheritance (14). The sharing of only one MHC haplotype or the sharing of no haplotypes is not due to incomplete penetrance but rather to the common occurrence of unaffected MHC susceptibility gene homozygous parents. Such parents usually do not have disease because they lack non-MHC genes in proper dosage or, even if they have them, the composite susceptibility genotypes may not be penetrant.

The second retrospective approach is to analyze the homozygote-heterozygote distribution of marker alleles in a population of unrelated but ethnically similar persons with the trait (patients) (15). If there is a marker that is common among patients and the susceptibility gene determines the trait in a recessive fashion, the distribution of homozygotes, heterozygotes and non-carriers of that
marker should fit the Hardy-Weinberg equilibrium. This method works perfectly if the marker is the susceptibility allele or if the marker allele is always on the same haplotype as the susceptibility allele in the population (shows absolute positive linkage disequilibrium or is part of a completely conserved extended haplotype or its fragment). This is largely true of the \textit{BF*F1} MHC marker for type 1 diabetes (16).

On the other hand, if the disease is dominantly determined, only the gene on one chromosome is needed and the frequency of alleles on the other chromosome will reflect the background for that ethnic group (15, 16). The distributions of HLA-B27 in ankylosing spondylitis (15) and HLA-DR4 in Jewish patients with pemphigus vulgaris (10) have been shown to fit dominant genetic determination based on such analyses. Because half of first-degree relatives of patients with pemphigus vulgaris have low levels of the pemphigus antibody, formal segregation and linkage analyses were possible and provided strong corroborative evidence for dominant class II region MHC-linked inheritance for this trait (17). Unfortunately, this ability to detect expression of the trait in asymptomatic carriers is unusual among autoimmune diseases. It would be helpful to have other approaches to the study of polygenic, MHC-controlled phenomena involving partially penetrant genes.

The prospective approach

Animal and in vitro models have been used to test prospectively whether one or another marker is a susceptibility or protective allele by transfection of cells \textit{in vitro} or rendering inbred strains transfenic for or devoid of the gene in question. These have proved to be fruitful in the exploration of the role of HLA-B27 in ankylosing spondylitis (18) and of HLA class I proteins as targets for natural killer (NK) cells (19).

In a second approach, described here, one can analyze for a suspected MHC-determined trait in individuals known to be homozygous, heterozygous or non-carriers of MHC susceptibility genes. In this respect, associated conserved extended haplotypes and their fragments are extremely useful and analogous to the MHC in syngenic and recombinant mice. Because of the fixity of DNA within the MHC of the extended haplotypes, all or nearly all independent instances of a trait-associated extended haplotype in a population (affected or not) carry the same susceptibility genes. By analyzing populations of persons homozygous, heterozygous or non-carrying for a known susceptibility extended haplotype for expression of a trait, one can decide if the trait is dominantly or recessively determined.

For example, it was concluded retrospectively from population studies of nonresponders to the HBsAg vaccine (20) that there was an increased frequency of persons (including homozygotes) with the extended haplotype [HLA-B8, SC01, DR3]. This suggested recessive determination of the nonresponse. In prospective studies, unrelated homozygotes and heterozygotes for [HLA-B8, SC01, DR3] were immunized with standard courses of HBsAg vaccine (21). The homozygotes were nonresponders whereas the heterozygotes produced anti-HBs, with no overlap in antibody levels. Family studies con-
firmed that the anti-HBs response was dominant and the failure to respond was recessive and strongly MHC linked (22). Thus, the HLA-B8, SC01, DR3) haplotype usually lacks an immune response gene for HBsAg. The response is dominant and the nonresponse is recessive.

In our proposed analysis, homozygotes for the susceptibility gene-bearing haplotype have two susceptibility genes, one on each chromosome, whereas heterozygotes have only one, if the background frequency of the susceptibility gene on all other haplotypes is low. If baseline penetrance is 100% and the trait is monogenic and recessive, all homozygotes, but no heterozygotes or non-carriers will express the trait (i.e. Mendelian inheritance will hold). If baseline penetrance is less than 100% but detectably above zero or if more than one gene is required for susceptibility, the observed lower prevalence of affected homozygotes reflects the population penetrance of the susceptible genotype (homozygotes and non-carriers will still be substantially unaffected). If the trait is dominantly expressed, both homozygotes and heterozygotes but few non-carriers will be affected. Again, incomplete baseline penetrance and the requirement for a number of non-MHC genes for trait expression will not change which group of subjects is affected.

Mechanisms of incomplete baseline penetrance: environmental or intrinsic?
Classically, the non-genetic (“environmental”) portion of incomplete penetrance has been defined by concordance rates for the trait in monozygotic twin pairs ascertained because one twin expresses the trait. Commonly invoked to explain this aspect of incomplete penetrance is some environmental factor, particularly viral infection. This explanation has often been extended to explain the lower apparent penetrance in MHC-identical sibs than in monozygotic twins by the unstated assumption that there is a larger difference in the environments of sibs than of twins. Because many of the diseases with incomplete penetrance are characteristically not manifest at birth, the notion of an environmental triggering event is attractive. However, recent studies of persons at high risk for type 1 diabetes provide strong support for the onset of abnormalities many years before overt disease develops (23, 24). Therefore, it is conceivable that some of these are, in fact, present at birth or earlier in individuals destined later to have frank disease.

Another reason for considering an environmental trigger for MHC-associated diseases with incomplete penetrance is the fact that most are autoimmune. The trigger may be exposure to a foreign antigen with high similarity to self antigens and thus capable of breaking immunologic tolerance to self antigens, such as to the β cells of the pancreas in type 1 diabetes. In our view, if there is an environmental trigger for autoimmune phenomena and disease, it affects genetically susceptible individuals to the same degree. This argument is strongly supported by the observation that dizygotic twins of patients with type 1 diabetes have precisely the same concordance rate for disease as sibs in general (25). In other words, the entire difference in rates of disease between sibs and monozygotic twins must be due to differences in susceptibility gene frequency. Differences in environmental triggers cannot be responsible for the difference (although environmental factors may, of course, affect the rate of baseline penetrance). The consequence of this view is that baseline penetrance acts in any genetic model as a mathematical multiplier that can be applied to all genetically susceptible individuals in the population, whether related or not. For example, if 10% of a specified group of subjects is genetically susceptible and baseline penetrance is 50%, 5% will be affected.

An alternative to an environmental trigger determining whether or not the trait will be manifest in a genetically susceptible person (Fig. 2A for the dominant situation, but also applicable to recessive traits) is an “intrinsic” event in a susceptibility gene, a kind of on-off switch (Fig. 2B, shown only for the dominant case). This would presumably be set during embryogenesis and would behave as a stochastic process, much as radioactive decay is intrinsic to certain isotopes. In this view, if the intrinsic process were influenced by environmental factors, the latter would affect all susceptibles in the same way and to the same extent.

There is an important conceptual difference between an environmental trigger and an intrinsic on-off switch basis for baseline penetrance. The trigger acts on the whole organism, but the intrinsic switch acts at the gene level. This is shown in Fig. 2 for the dominant case. Implicit in the trigger mechanism is that a genetically susceptible person is converted to expression of the trait by exposure to something such as a pathogen, an immunogen or an allergen. In dominant expression of susceptibility genes, the two mechanisms predict different apparent penetrances in homozygotes and heterozygotes. For the external triggering mechanism (Fig. 2A), the apparent penetrance per person in homozygotes and heterozygotes for marker susceptibility genes should be the same, since the environmental effect should trigger trait expression whether the subject has one or two susceptibility alleles. The ratio of affected homozygotes to affected heterozygotes will remain 1 no matter what the baseline penetrance or the requirement for non-MHC susceptibility genes.

If penetrance is externally triggered and genetic expression is recessive, any person with two copies of the susceptibility gene will express the trait (if the other susceptibility genes are present in the correct number) if properly triggered by the external agent or event. In intrinsic penetrance of a recessive susceptibility gene, only homo-
### A. DOMINANT EXPRESSION, EXTRINSIC TRIGGER

- **3 AFFECTEDS**
  - TRIGGER
  - TRIGGER
  - TRIGGER

- **HOMOZYGOTES**: DD, DD, DD, DD, DD, DD
- **HETEROZYGOTES**: ND, ND, ND, ND, ND, ND

- **PENETRATION**
  - 0.5 PER PERSON
  - 0.3 PER PERSON

- **3 AFFECTEDS**

### B. DOMINANT EXPRESSION, INTRINSIC SWITCH

- **5 AFFECTEDS**
  - TRIGGER
  - TRIGGER
  - TRIGGER

- **HOMOZYGOTES**: DD, DD, DD, DD, DD, DD
- **HETEROZYGOTES**: ND, ND, ND, ND, ND, ND

- **PENETRATION**
  - 0.2 PER HAPLOTYPE
  - 0.3 PER HAPLOTYPE

- **3 AFFECTEDS**

**Fig. 2.** Diagrammatic representation of extrinsic and intrinsic expression of a partially penetrant dominant gene D in homozygous (DD) and heterozygous (ND) individuals, shown as circles. Extrinsic expression (A) is shown as an arrow mediated by a trigger and penetrance is a rate per person. Intrinsic expression (B) is shown by an activated (blackened) gene and an arrow. In this model, the rate of expression is per gene or haplotype, with a person manifesting the trait if one or both susceptibility genes are “activated.” Note that for the extrinsic model, the ratio of affected homozygotes to affected heterozygotes is 1 (3:3 in this example). In contrast, in the intrinsic model there are more affected homozygotes than affected heterozygotes (5:3 in the example given).

zygous individuals with activation of the gene on both chromosomes will manifest the trait. Thus, both mechanisms produce the same end result and there is no obvious way to distinguish which mechanism is operative.

On the other hand, if the susceptibility gene is dominant and penetrance is intrinsic (Fig. 2B), the rate of trait expression in homozygotes and heterozygotes will differ, since the homozygote with two susceptibility genes has a greater chance (up to two-fold) of the intrinsic switch being turned on. For dominant genetic determination, trait expression (apparent penetration or p) in susceptibility haplotype heterozygotes is equivalent to prevalence of the trait in this group of subjects. This follows from the fact that heterozygotes have only one susceptibility gene that is either on or off. The per person prevalence is the same as the per chromosome apparent penetrance. In haplotype homozygotes, the situation is more complex. Homozygotes with both genes turned on, as well as those with only one gene turned on, may express the trait. For an apparent intrinsic penetrance p, the frequency of affected homozygotes will be $p^2 + 2p(1-p)$ (or affected homozygotes with two expressed genes $(p^2)$ + affected homozygotes with only one expressed gene $2p(1-p)$).

For example, at an apparent penetrance of 0.2, determined from trait expression in haplotype heterozygotes, haplotype homozygotes with both susceptibility genes expressed will constitute 0.2×0.2, or 0.04, of all homozygotes studied. Haplotype homozygotes with only one expressed gene, on the other hand, will constitute 2p(1–p) or, in our example, 0.2 (expressed)×0.8 (non-expressed)×2 (since either chromosome can be expressed) or 0.32 of all homozygotes studied. Thus, for a penetrance of 0.2, 0.36 (0.04+0.32) of haplotype homozygotes will be affected and the ratio of affected homozygotes to affected heterozygotes will be 0.36/0.20 or 1.80.

Another and mathematically equivalent way of looking at the same problem is to note that penetrance in the homozygote with two susceptibility genes is ideally $2\times p$. This needs to be “corrected” by subtracting those individuals ($p^2$) in whom both genes are activated ($2p–p^2$). $2p−p^2$ is simply a rearranged form of $p^2 + 2p(1–p)$.

**Fig. 3** shows the ratio of affected homozygotes to heterozygotes in relation to intrinsic penetrance. If baseline penetrance is 100%, the ratio is 1, Mendelian inheritance holds and all homozygotes and heterozygotes with susceptibility genes are affected. In the dominant intrinsic model, the process is stochastic and, at low penetrance, the homozygote with two susceptibility alleles has nearly twice the opportunity for expression and therefore should be nearly

**Fig. 3.** The ratio of affected homozygotes to heterozygotes for a partially penetrant dominant susceptibility gene in relation to penetrance based on the intrinsic mechanism. Note that at very low penetrance, the ratio of affected homozygotes to affected heterozygotes approaches 2, whereas if penetrance is 1, the ratio is 1 (Mendelian inheritance).
twice as often affected as the heterozygote. The requirement for non-MHC genes will have no affect on this ratio since their frequency should be the same in randomly selected MHC homozygotes, heterozygotes and non-carriers. Affected heterozygotes, of course, have only one expressed susceptibility gene and the rate at which they are affected equals the apparent penetrance rate in this group of subjects. Thus, a test of the mechanism for penetrance is the ratio of the fraction of prospectively ascertained affected homozygotes for a dominant trait-associated extended haplotype to that of affected heterozygotes.

For recessive expression, if \( p \) = decimal fraction of all susceptibility genes that are expressed (as above), and \( 1-p \) = the fraction of non-expressed susceptibility genes, \( p^2 + 2p(1-p) \) describes the frequency of subjects with “activated” susceptibility genes. However, since trait expression can only occur when both chromosomes have activated genes, the term \( 2p(1-p) \) drops out because subjects with only one activated gene do not express the trait. Apparent penetrance is the square root of the frequency of the trait in haplotype homozygotes. Thus, without more information, the external trigger and intrinsic mechanisms for recessive gene expression are indistinguishable.

Our model proposes a simple and precise mechanism for higher penetrance in dominant susceptibility gene homozygotes than heterozygotes. It should be recognized that the actual mode of genetic expression may be more complicated than that outlined. Although other mechanisms might also predict higher expression in homozygotes, we prefer the simpler explanation.

Applications

When this analysis was applied to immunoglobulin deficiencies determined by the conserved extended haplotype [HLA-B8, SC01, DR3] (26), IgA deficiency and IgG4 deficiency were found essentially only in homozygotes, suggesting recessive inheritance. On the other hand, IgD and IgG3 deficiency occurred in both homozygotes and heterozygotes at apparent penetrances of 0.37 and 0.20 for IgD and 0.30 and 0.17 for IgG3 deficiency, consistent with dominant inheritance. Ratios of affected homozygotes to affected heterozygotes were thus 1.8 and 1.7, respectively, remarkably close to those predicted by the formula given above for intrinsic penetrance.

In a study of Basques and [HLA-B18, F1C30, DR3] (27), the only deficiency was in IgD present in unrelated homozygotes and heterozygotes with apparent penetrances of 0.37 and 0.19, for a ratio of 1.9. This suggests that, for these traits, the intrinsic model of baseline penetrance applies. As stated previously, since subjects were grouped solely on the basis of homozygosity, heterozygosity or non-carrying of [HLA-B8, SC01, DR3] or [HLA-B18, F1C30, DR3], it is reasonable to assume that the frequencies of non-MHC susceptibility genes are the same in all three groups. Therefore, baseline penetrance is probably intrinsic for the MHC susceptibility genes on these haplotypes.

The prospective genetic analysis of MHC-determined traits outlined here is powerful and, if the trait is dominantly expressed, can provide evidence for the mechanism of baseline penetrance. It has an additional advantage in that, by studying subjects of specified genetic makeup, the relationship between phenotype heterogeneity and specific susceptibility haplotypes can be explored. Although the analysis is only helpful in testing the mechanism of baseline penetrance in the dominant situation, it may be that the intrinsic mechanism is operative in MHC-recessive diseases, such as type 1 diabetes. This has important implications for unraveling the identity, nature, and function of MHC genes involved in susceptibility to autoimmune disease, including type 1 diabetes. Moreover, by the prospective selection of individuals with informative fragments of susceptibility gene-bearing extended haplotypes that presumably arose from ancient crossovers, localization of these genes can be greatly facilitated. In work in progress, this fragment approach is being used to localize MHC susceptibility genes for the immunoglobulin deficiencies.

References

