

The Haplotype Structure of the Human Major Histocompatibility Complex

Chester A. Alper, Charles E. Larsen,
Devendra P. Dubey, Zuheir L. Awdeh,
Dolores A. Fici, and Edmond J. Yunis

ABSTRACT: There is great interest in the use of single-nucleotide polymorphisms (SNPs) and linkage disequilibrium (LD) analysis to localize human disease genes. The results suggest that the human genome, including the major histocompatibility complex (MHC), consists largely of 5- to 200-kb blocks of sequence fixity between which random recombination occurs. Direct determination of MHC haplotypes from family studies also demonstrates similar-sized blocks, but otherwise gives a very different picture, with a third to a half of Caucasian haplotypes fixed from *HLA-B* to *HLA-DR/DQ* (at least 1 Mb) as conserved extended haplotypes (CEHs), some of which encompass more than 3 Mb. These fixed haplotypes differ in frequency both in different Caucasian subpopulations and in Caucasian patients with HLA-associated diseases, complicating disease susceptibility gene localization. The inherent inability of LD analysis to “see” DNA fixity beyond three markers contributes to the failure of

SNP/LD analysis to define in detail or even detect CEHs in the MHC and probably elsewhere in the genome. More importantly, the use of statistical analysis, rather than direct haplotype determination and counting, fails to reveal the details of haplotype structure essential for gene localization. Given the oversimplified picture of the MHC (and probably the rest of the genome) provided only by SNP/LD-defined blocks, it is questionable whether this approach will be of great help in disease susceptibility gene localization or identification. *Human Immunology* 67, 73–84 (2006). © American Society for Histocompatibility and Immunogenetics, 2006. Published by Elsevier Inc.

KEYWORDS: major histocompatibility complex; susceptibility genes; complex diseases; single-nucleotide polymorphisms; linkage disequilibrium

ABBREVIATIONS

CEH conserved extended haplotype
HBsAg hepatitis B surface antigen
HLA human leukocyte antigen
LD linkage disequilibrium
MHC major histocompatibility complex

PCR polymerase chain reaction
PV pemphigus vulgaris
RFLP restriction fragment length polymorphism
SNP single-nucleotide polymorphism
T1D type 1 diabetes

INTRODUCTION

There has been considerable interest in the use of single-nucleotide polymorphisms (SNPs) to map and identify human disease genes, particularly those determining susceptibility to complex diseases [1–3]. These SNP polymorphisms occur, on average, every 290 bp of genomic

DNA [4] and form haplotypes that are often deduced from studies of the DNA of many individuals by linkage disequilibrium (LD) analysis. These haplotypes identify islands of nucleotide sequence stability (blocks) between which meiotic crossing over occurs (and occurred historically) as “hotspots.” The blocks are operationally defined by strong LD (D' near 1) between the outermost defining SNP pairs and a greater than 19-fold ratio of strong to weak LD (D' significantly < 1) among all internal pairwise comparisons [2, 3]. With this definition, these studies have determined that the vast majority of such blocks are between 5 and about 200 kb in size [1], although the possibility of occasional, larger blocks (up to 804 kb) has also been suggested [5]. It is essential to

From The CBR Institute for Biomedical Research, Boston, MA 02115, USA (C.A.A., C.E.L., D.P.D., Z.L.A., D.A.F., E.J.Y.); Departments of Pediatrics (C.A.A.), Medicine (C.E.L.), and Pathology (E.J.Y.), Harvard Medical School, Boston, MA 02115, USA; and Dana-Farber Cancer Institute, Boston, MA 02115, USA (E.J.Y.).

Correspondence should be addressed to: Chester A. Alper, M.D., The CBR Institute for Biomedical Research, 800 Huntington Avenue, Boston, MA 02115, USA; Tel: (617) 278-3333; Fax: (617) 278-3493; E-mail: alper@cbr.med.harvard.edu.

note that there has been much debate in the literature with respect to the best criteria for establishing blocks. Not stressed, but sometimes found, is LD between blocks that may define stretches of fixity of up to 500 kb on 38% of chromosomes [1]. Remarkably, a relatively small number (two to four) of SNP haplotype variants account for 80–95% of all observed chromosomes [1] and are found in all human populations studied, although relative frequencies vary [2].

The human major histocompatibility complex (MHC) on chromosome 6 has been studied intensively. It is of great interest because it contains highly polymorphic genes that determine immune function, susceptibility to a remarkably large number of complex diseases, and the outcome of tissue transplantation. We and others have used pedigree analysis and a direct observational approach, an extension of the counting methods proposed by Smith and co-workers [6, 7], to determine population MHC haplotype composition from family studies. One simply counts the number of observed similar (identical or near-identical for HLA-B, complotype, and HLA-DR/DQ) haplotypes (conserved extended haplotypes or CEHs) in any specified group of families to define differences in MHC haplotype frequencies in populations that differ in ethnicity or in disease state (association studies), or to determine whether DNA conservation extends to intervening or nearby polymorphic loci. In the case of a disease, the haplotype composition of chromosomes occurring in patients in the families is compared with the composition of all other chromosomes occurring in those families, thus providing family control haplotypes [8, 9]. For these purposes, as well as for showing that the fixity of these similar haplotypes is statistically significant by LD analysis, simple χ^2 statistics suffice [10, 11].

The present article describes pedigree-determined HLA specificity- and allele-defined MHC haplotypes in Caucasian subjects of several different ethnicities and those in one particular Caucasian ethnic subgroup and in patients with two different MHC allele-associated diseases. The conclusions drawn from these studies are then compared with those based on the SNP/LD-defined structure of the human genome, including such studies in the MHC.

METHODS

Subjects

Nuclear families of normal Caucasian subjects, or of patients with type 1 diabetes (T1D), pemphigus vulgaris, gluten-sensitive enteropathy, ragweed allergy, dermatitis herpetiformis, multiple sclerosis, or congenital adrenal hyperplasia, or of patients who were HBsAg vaccine nonresponders or had other immune or autoim-

mune disorders were studied and reported previously [8, 12–20]. Disease haplotypes were those occurring in a patient. There were 353 T1D haplotypes, as well as 61 Ashkenazi Jewish pemphigus vulgaris (PV) patient haplotypes. Normal family control haplotypes were those not occurring in a patient [8, 9, 21] and totaled 2000 normal Caucasian haplotypes (including a subpopulation of 195 normal Ashkenazi Jewish haplotypes). All MHC haplotypes were defined by their HLA-A, HLA-B, complotype, and HLA-DR specificities. PV is a blistering autoimmune disease known to be associated with HLA-DR4 (DRB1*0402) in Ashkenazi Jews [14].

LD Analysis

For the analysis of two-point, three-point, four-point, and five-point LD, 86 pedigree-determined haplotypes for which high-resolution *HLA-DRB1*, *-DQA1*, and *-DQB1* alleles, as well as HLA-A, HLA-B, and complotype specificities, were available, were randomly paired to create 43 “people.” We did this so that the data we analyzed would correspond to unphased data from people as normally used for LD analysis. Allelic LD was defined as $D = P_{ij} - (p_i \cdot q_j)$, where P_{ij} is the frequency of the gamete carrying the i th allele of one locus and the j th allele of another locus, and p_i and q_j are the frequencies of the i th and j th alleles of the two loci [22, 23]. To calculate LD involving two, three, four, and five loci (D), we used a standard method [24]. Normalized LD (D') was calculated as suggested by Lewontin [25].

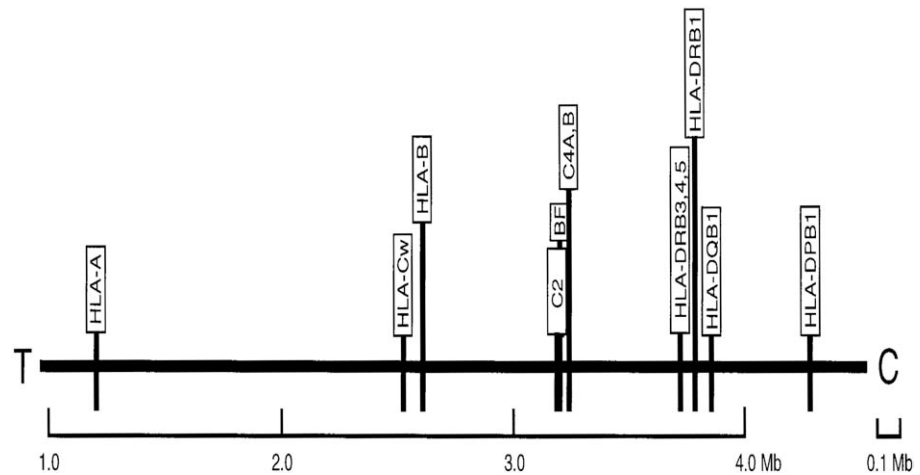
HLA Typing

Peripheral blood mononuclear cells of patients, their parents, and their siblings were analyzed for HLA-A, -B, and -DR specificities [26] and haplotypes were determined by segregation analysis. For some haplotypes, HLA-A, HLA-B, HLA-Cw, HLA-DRB1, HLA-DQA1, and HLA-DQB1 alleles were determined by polymerase chain reaction (PCR) amplification of genomic DNA and dot-blot analysis using sequence-specific oligonucleotide probes [27] or by PCR and sequence-specific primers [28, 29]. In all instances, specific allele assignments to CEHs corresponded to those reported [30]. MHC complement haplotypes (complotypes) were determined as described previously [31].

RESULTS AND DISCUSSION

The MHC comprises more than 3 Mb of genomic DNA on human chromosome 6p21.3. Figure 1 illustrates only a few of the genetic loci of this region: *HLA-A*, *HLA-Cw*, *HLA-B*, *C2*, *BF*, *C4A*, *C4B*, *HLA-DRB3*, *-4*, and *-5*, *HLA-DRB1*, *HLA-DQB1*, and *HLA-DPB1*. Studies over the past several decades of genes within the human MHC have yielded a picture of polymorphism and genomic structure that in some ways coincides with that

FIGURE 1 Map of the MHC indicating a few of its genes. Distances are drawn to scale (see legend), but these may vary at many locations in different haplotypes as a result of limited polymorphic DNA insertions, deletions, or gene duplications. Gene locations were taken from the Sanger Institute MHC consensus gene list (<http://www.sanger.ac.uk/HGP/Chr6>), and the distance (in megabases, Mb) from an arbitrary point telomeric (T) to *HLA-A* to the region centromeric (C) to *HLA*DPB1* is indicated.



generated from the more recent studies using SNPs and yet is far more diverse and complex.

The point of agreement between the two approaches lies in the occurrence in the MHC of small fixed blocks of DNA around which meiotic recombination occurs and has occurred historically. The first of these to be described was the *complotype* [32], a region of approximately 75–120 kb (its size varies among individual chromosomes) designated by its *BF*, *C2*, *C4A*, and *C4B* alleles in that (arbitrary) order. The complotype SC01, for example, contains *BF*S*, *C2*C*, *C4A*Q0*, *C4B*1*. Other blocks in the MHC comprise alleles of the closely situated genes for *HLA-Cw* and *-B*, as well as those for *HLA-DR* and *-DQ* (including *DRA*, the various *DRB* loci, *DQA1*, and *DQB1*). It should be pointed out that the MHC allele-defined blocks are far more polymorphic than the SNP haplotype-defined blocks. For example, in Caucasians there are 14 protein- and RFLP-defined complotypes [33] with a frequency of 0.02 or higher. Formal LD analysis derived from the study of MHC alleles, microsatellites, and SNPs [30, 34] in populations of unrelated individuals supports the presence of *HLA-Cw/B* and *HLA-DR/DQ* blocks, but, surprisingly, complotypes were not detected [34].

At least a third of common (frequency > 0.01) control European Caucasian MHC haplotypes are fixed from *HLA-B* through the complotype to *HLA-DR/DQ* [10, 30], a distance of more than 1 Mb. If one includes rarer *HLA-B*, complotype, and *HLA-DR/DQ* haplotypes (at least 5 independent examples in the 2000 control haplotypes (frequency > 0.0025)) presented here, fixed haplotypes account for a little more than half of all Caucasian MHC haplotypes. Moreover, these conserved extended haplotypes (CEHs) [10] or, as subsequently designated, ancestral haplotypes [35], occur at widely differing frequencies not only in the major human ethnic groups, but even within subsets of those populations. Furthermore,

virtually all of the MHC allele–disease associations involve marker alleles of CEHs.

The essential quality of a CEH is the fixity of its DNA over at least 1 Mb of MHC DNA. Historically, this has been defined by serologic *HLA-B*, complotype, *HLA-DR*, and *HLA-DQ* specificities [10]. With technological refinements, it is clear that, if a specificity is determined by multiple DNA-defined alleles, CEHs carry only one (or, unusually, at the most a few) of these alleles [36]. Moreover, intervening DNA manifests identical or near-identical sequences [37–39, for example] without apparent recombination. In addition, different *HLA-A* alleles are usually relatively fixed on specific CEHs. There is little difficulty in applying these criteria in defining CEHs with a frequency ≥ 0.01 in a population. Our population of 2000 normal Caucasian haplotypes has 15 such common CEHs with a minimum of 20 independent instances each. The problem arises with the lower-frequency CEHs (31 possible CEHs with frequencies of 0.0025–0.0095). These clearly need to be studied in more detail before one can be certain of the fixity of their DNA, although almost all have restricted *HLA-A* variation. Finally, the ability to identify CEHs depends heavily on the study population. For example, the CEH [*HLA-B55*, *SB45*, *DR6*] occurs only twice in our 2000 normal Caucasian haplotypes (0.001), yet has a frequency of 0.12 among non-Jewish patients with PV [17].

Because the MHC contains many genes affecting immune function and some individual CEHs may extend over 3 Mb from at least *HLA-A* to *HLA-DPB1* (around 0.1% of the human genome), it is not surprising that CEHs provide genetic markers for susceptibility for a large number of (mostly autoimmune) diseases. For example, in studies of Caucasian patients with T1D, it was initially noted that *HLA-B8*, *-B18*, and *-B15(62)* were markedly increased in frequency [40, 41]. Later, it was demonstrated that *HLA-DR3* and *-DR4* were strongly

associated [42, 43]. Still later, the complement alleles BF*F1, C2*B, and C4B*3 were found to be associated. All of these markers belong to collections of MHC specificities/alleles (at that time not considered to be necessarily haplotypes) referred to as *supratypes* [44]. In family studies of T1D patients, it was demonstrated that, in fact, all of these allotypes are markers for CEHs [8] or ancestral haplotypes [45].

In all instances in which intervening DNA on independent examples of CEHs has been studied, it has been identical or has manifested highly restricted variation [10, 37–39, 46–52], including SNPs [53, 54]. Wherever analyzed, each CEH carries a specific allele (or, at most, a few closely related specific alleles) determined from DNA [30]. As an extreme example, the [HLA-B38, SC21, DRB1*0402] CEH exhibits limited variation in many of its recognized elements: HLA-B38 may be HLA-B35, and SC21 may be SC31 or (rarely) SC20. Alternatively, there may be three closely related CEHs: [HLA-B38, SC21, DR4], [HLA-B38, SC31, DR4], and [HLA-B35, SC31, DR4]. Variability on CEHs is often greatest in microsatellite loci, which are both hypermutable, owing to “slippage,” and more susceptible to incorrect allele assignment than standard MHC alleles. Functional identity of specific CEH homozygotes, as in the lack of the mixed lymphocyte reaction [55, 56], the lack of an immune response to HBsAg [15], or MHC-determined serum immunoglobulin deficiencies [20, 57–59], presumably reflects structural identity of the MHC susceptibility genes involved that are within the fixed DNA region. For more details about CEHs, see a recent review [30].

To illustrate dramatically the remarkable diversity of human MHC haplotypes and to demonstrate how this varies sharply among patients with different MHC-associated diseases and different subethnicities, consider HLA-DR2-, -DR3- and -DR4-carrying MHC haplotypes directly observed in Caucasian family studies in Boston (Figure 2). Table 1 provides the key to Figure 2 and outlines the composition of CEHs with a frequency of ≥ 0.02 in any of the four populations described earlier. HLA-A specificities occurring with a frequency > 0.22 (arbitrary cutoff) on individual instances in the normal population of each of these haplotypes, as well as further allele and specificity data, are also provided.

A conclusion easily drawn from Figure 2 is that specific CEHs (solid color from HLA-B to HLA-DR representing at least 1 Mb of genomic DNA) and their fragments (smaller, similarly colored segments) differ markedly in frequency on DR2, DR3, and DR4 haplotypes in all the study groups. Thus, there is clear evidence for the common but variable fixity of genomic DNA well beyond the usual 200 kb or so maximum for SNP/LD-detected blocks. A second obvious point is that these high-frequency CEHs and their fragments constitute a large part of all of the

Caucasian haplotype groups studied (including the approximately 50% of normal Caucasian MHC haplotypes with HLA-DR1, -DR5, -DR6, -DR7, or other HLA-DRs, not shown). It is also clear that there is variation in the number of common CEHs carried by different HLA-DR haplotypes, with HLA-DR4 having the most diversity, and HLA-DR3, the least.

The differential diversity of CEHs and their restricted diversity among ethnic subgroups such as Ashkenazi Jews compared with general Caucasians are also obvious. The [HLA-B38, SC21, DR4] CEH (including its variations), for example, has a frequency of 0.133 on normal Ashkenazi haplotypes, 0.543 on Ashkenazi PV haplotypes, and 0.012 on general Boston Caucasian haplotypes. Most of the last were from Ashkenazi controls. It is of interest that [HLA-B38, SC21, DR4] has an increased frequency among T1D haplotypes [8], contributed mostly, again, by Jewish patients. This suggests that this CEH carries susceptibility genes for both T1D and PV. In contrast, [HLA-B8, SC01, DR3] has frequencies of 0.09 among general control haplotypes, 0.04 among control Ashkenazi, and 0.18 among T1D haplotypes, and was not found at all among the Ashkenazi PV haplotypes. Thus, this CEH carries a susceptibility gene for T1D but not PV. The CEH [HLA-B8, SC01, DR3] is most common in the British Isles and has its highest normal frequency (0.22) among our Irish haplotypes (data not shown). It also carries susceptibility for a remarkable array of autoimmune diseases.

The well-known increased frequency of HLA-DR3 and -DR4 markers in our T1D patients is accompanied by frequency increases in specific CEHs: [HLA-B8, SC01, DR3], [HLA-B18, F1C30, DR3], [HLA-B62, SC33, DR4], [HLA-B60, SB42, DR4], among others (Figure 2). The specificities in bold type include the previously noted markers for T1D. Most (~70%) of the DR4 haplotypes are DRB1*0401 and carry HLA-DQB1*0302. The paucity of HLA-DR2 haplotypes in T1D patients seen in Figure 2 involves a proportional marked decrease in the frequency of the protective CEH [HLA-B7, SC31, DR2]. From these observations we can conclude that those CEHs that are increased in frequency in patients, even when they occur in the general population, carry genes for susceptibility to T1D, whereas those that are decreased carry “protective” alleles. Whether or not different CEHs carry different true susceptibility alleles is not known, but it is certainly not the case in this common complex disease that there is only one set of haplotype markers for susceptibility. Similarly, the two CEHs that mark susceptibility to PV in Europeans have largely different marker alleles.

It is further evident from Figure 2 that the DNA fixity of CEHs often extends through HLA-A. This is manifested by the common (> 0.22) occurrence of one or

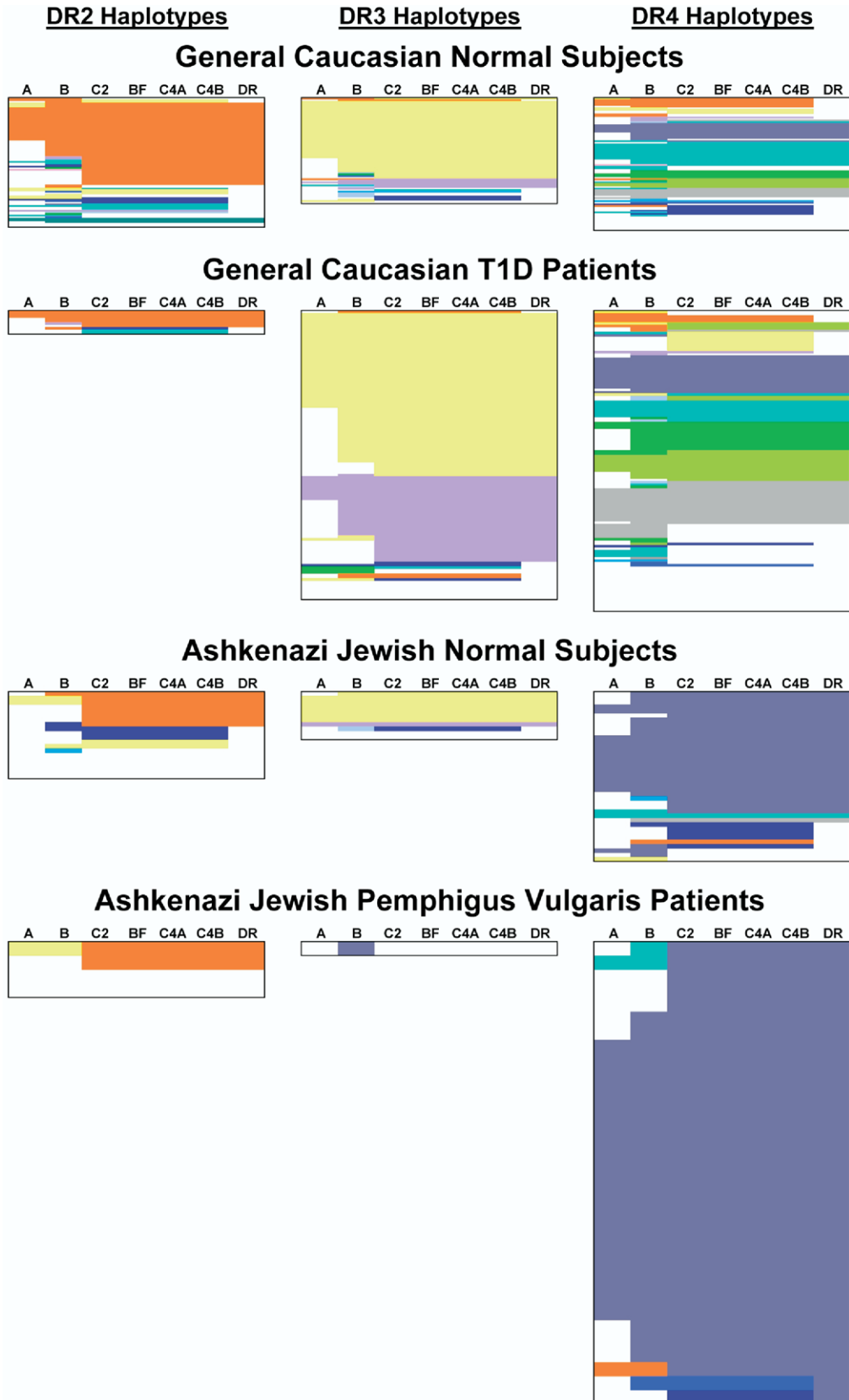


FIGURE 2 Directly determined MHC haplotypes carrying HLA-DR2, -DR3, and -DR4, showing their HLA-A, HLA-B, complo-type, and HLA-DR specificities. The four populations of haplotypes are: normal Caucasian control haplotypes ($n = 2000$), Caucasian T1D ($n = 353$), Ashkenazi control ($n = 195$), and Ashkenazi PV ($n = 61$). The ordinate for each haplotype population is the decimal fraction of the total population of those haplotypes. The key to the color coding and more information on the CEHs involved are given in Table 1. Because some HLA-B specificities (*e.g.*, B44, B35, B62) are, at this level of typing, shared by two or more CEHs, we have assigned these freestanding specificities to a CEH on the basis of the HLA-A or -C allele (not shown) with which they occur on that particular haplotype [30, 49, 52]. The complo-type SC31 is on around 40% of normal Caucasian MHC haplotypes and is shared by several CEHs. It can be subdivided at the molecular level, with each CEH with SC31 carrying only one of three subsets of DNA restriction fragment length polymorphism (RFLP) markers [33], but this has not been done for most of the haplotypes shown. SC31 is considered only part of the respective CEH when it occurs with the appropriate HLA-B or HLA-DR specificity or allele of that CEH. Colors shown here but not included in Table 1 represent fragments of CEHs bearing HLA-DR1, -DR5, -DR6, or -DR7 or lower frequency CEHs.

TABLE 1 HLA-DR2-, DR3-, or DR4-containing CEHs with a frequency of 0.02 or higher in at least one of the specified (Fig. 2) populations of Caucasian subjects

HLA-A ^a	CEH	Gen Cauc ^b	T1D Gen Cauc	Ashk Jewish	PV Ashk Jewish	HLA alleles
A3, A2	B7, SC31, DR2	0.069	0.015	0.005	—	HLA-A*0301/0201, B*0702, DRB1*1501
A1	B8, SC01, DR3	0.086	0.183	0.036	—	HLA-A*0101, B*0801, DRB1*0301
A30	B18, F1C30, DR3	0.006	0.077	0.005	—	HLA-A*3001, B*1801, DRB1*0301
A26	B38, SC21, DR4	0.012	0.041	0.133	0.543	HLA-A*2601, B*3801, DRB1*0402
A2	B44, SC30, DR4	0.026	0.020	0.010	—	HLA-A*0201, B*4402, DRB1*0401
A2	B62, SC33, DR4	0.010	0.044	0.005	—	HLA-A*0201, B*1501, DRB1*0401
A31	B60, SC31, DR4	0.009	0.035	—	—	HLA-A*3101, B*4001, DRB1*0401
A2	B62, SB42, DR4	0.005	0.029	—	—	HLA-A*0201, B*1501, DRB1*0401

^aHLA-A specificities shown are those that have a frequency of at least 0.22 on independent examples of the indicated CEHs among these normal haplotypes. See Yunis et al. [30] for other alleles of these CEHs.

^bFrequency of each CEH in 2000 normal family control Caucasian haplotypes. Gen Cauc, general Caucasian; Ashk, Ashkenazi.

two HLA-A specificities characteristic of each CEH. Thus, 82% of our normal [HLA-B8, SC01, DR3] CEHs carry HLA-A1, 43% of [HLA-B38, SC21, DR4] carry HLA-A26, and about 65% of [HLA-B7, SC31, DR2] carry either HLA-A2 or -A3. This HLA-A,-B fixity is preserved in many of the HLA-A,-B fragments of CEHs (Figure 2). Even minor frequency HLA-A specificities on CEHs are characteristic of specific CEHs and probably reflect subpopulation differences (unpublished observations). Also not shown is the fact that fixity on some CEHs extends at least from HLA-A through at least HLA-DP [60, 61], a distance of 3.2 Mb.

Because virtually all the MHC markers of common complex diseases are components of fairly high-frequency CEHs [8, 62–65], there is a problem in localizing MHC susceptibility genes. As one can see from Figure 2, there is little specific enrichment of CEH fragments in patients with T1D that can be used in susceptibility gene localization. In fact, no MHC susceptibility gene for complex disease has been definitively identified [66]. Thus, the large regions of fixity on marker CEHs may be an impediment, rather than an aid, to susceptibility gene localization. On the other hand, fragment analysis has

been useful in roughly localizing susceptibility genes for PV [14, 17], asthma in ragweed pollen allergy [18], IgA deficiency [51, 67–69], myasthenia gravis [65], and gluten-sensitive enteropathy and dermatitis herpetiformis [19].

Whether fragment analysis is useful in susceptibility gene localization also depends partly on whether there is only one marker CEH in a patient population and whether this is relatively infrequent in the background population. Because it is usually the whole CEH rather than its individual alleles that marks susceptibility to T1D, it is misleading to use the odds ratios of any of the component alleles as guides to susceptibility gene localization or identification. Thus, in our T1D data, the allele with the highest odds ratio is BF*F1 at 8.5 (HLA-DRB1*0301, HLA-DQB1*0201 and HLA-DRB1*04, DQB1*0302 have odds ratios of around 4.5 each) [66]. Neither we nor anyone else has claimed that this means that the MHC T1D susceptibility gene is or is closest to BF.

Recently, there have been statistical approaches to assessing fixity in the MHC. In a study involving markers at 14 MHC loci in 50 CEPH families [34], no LD

TABLE 2 Frequencies of HLA-A-B and HLA-B-DR LD pairs in relation to frequencies of selected CEHs carrying HLA-DR2, -DR3, or -DR4^a

LD pairs					Family study and counting			
HLA-A	HLA-B	NI HF	HLA-B	HLA-DR	NI HF	CEH	NI HF ^b	HLA-A dec fract
A3	B7	0.028	B7	DR2	0.046	[HLA-B7, SC31, DR2]	0.069	A3: 0.37
A1	B8	0.064	B8	DR3	0.070	[HLA-B8, SC01, DR3]	0.086	A1: 0.82
A30	B18	0.017	B18	DR3	0.018	[HLA-B18, F1C30, DR3]	0.006	A30: 0.36
A2	B44	0.065	B44	DR4	0.013	[HLA-B44, SC30, DR4]	0.025	A2: 0.76
A26	B38	0.019	B38	DR4	0.041	[HLA-B38, SC21, DR4]	0.012	A26: 0.68

^a Adapted from Bodmer and Bodmer [74] and compared with our CEH dataset. The population studied was in London, UK. The family studies were in Boston.

^b NI HF, normal haplotype frequency (of 2000 normal haplotypes in our dataset); dec fract, decimal fraction of independent examples of the specified CEH among the 2000 normal haplotypes.

blocks beyond *HLA-C/B* and *HLA-DR/DQ* were found, not even among the complement genes, for which there are no well-documented crossovers. In another study, “extended haplotype homozygosity” analysis [70] used family studies and many markers, but not counting. Their data suggest that small blocks found in the MHC do not differ significantly in structure from the rest of the human genome. However, the results of haplotype determination were not particularly fruitful. Only a small fragment, at best, of the [HLA-B8, SC01, DR3] CEH, the most common CEH in northwestern European populations, was detected. Although the CEH [HLA-B7, SC31, DR2] was recognized, no other common CEHs were detected. On the other hand, haplotype-specific LD (*i.e.*, presupposing the existence of some, but clearly not all, specific CEHs) was detected when *HLA-B* was used as an anchor in a study of more than 600 unrelated subjects in the United Kingdom [11]. Haplotypes were deduced as “surrogate extended haplotypes.” Some such assignments did not correspond to haplotypes that occurred commonly or even at all among our 2000 normal Boston haplotypes (there were no instances of HLA-B*1402, DRB1*1101 or A*2501, B*1801, DRB1*0301) and a number of the tumor necrosis factor SNPs were incorrectly assigned [53]. Because the method depends on prior knowledge of CEHs [10] and is subject to errors in allele assignments to CEHs, it is of limited utility.

What is the reason for the discrepancy between the rich and complex but directly determined view of a portion of the human genome based on family studies and the rather simple picture of similar SNP haplotypes with similar frequencies in all populations as small blocks separating regions of diversity generated by frequent meiotic crossing over? One possible explanation is that the MHC is completely atypical of the rest of the genome. This cannot be the reason, because both SNP haplotype analysis and MHC allele multiple-locus LD analysis produce the same small block results as in the rest of the genome [34, 70]. Moreover, the MHC has been studied far more intensively than the rest of the

genome. It may, in fact, be considerably less atypical than some might imagine. Furthermore, SNP analysis of human sperm has revealed the same kinds of blocks separating regions of frequent crossing over in the MHC as elsewhere in the genome [71, 72]. However, this is not at all the same as assessing the extent of fixity of DNA at the population level. We have dealt elsewhere with discrepancies between the hotspots detected in sperm meioses and those postulated from ancient recombinations by haplotype analysis [36]. In any case, in our view, it is likely that the large stretches of fixity of CEHs reflect the recent explosion of human populations [73] and, perhaps, population bottlenecks.

One might reasonably ask, given the high frequencies of CEHs, why these have not been detected by LD analysis. The curious answer is that they simply have been ignored. A quarter-century ago, two-point and three-point LD analysis first produced suggestive evidence that was not followed up. Table 2 outlines some HLA-A, -B and HLA-B, -DR specificity pairs exhibiting significant positive LD published previously [74]. Note that the pairs shown are parts of CEHs or predominant HLA-A markers for some of the CEHs presented in Table 1. They were not interpreted as DNA fixity at a time when natural selection was considered the likely basis for LD [75] and population stratification was also considered likely. “Young mutations,” migration, and inbreeding [74] were considered other possible but unlikely explanations. Three point (HLA-A, -B, -DR) LD analysis revealed the presence of population-characteristic significant combinations [76]. Somehow, this previous literature and the considerable MHC CEH literature (the references in this article are illustrative, not exhaustive) have been largely avoided by the investigators who analyze SNP/LD throughout the genome.

Why have these larger stretches of DNA fixity been ignored by many investigators? To help clarify the problem, consider LD analyses for the components of [HLA-B8, SC01, DR3] among the 43 control “people.” The frequency (by pedigree analysis) of the [HLA-B8, SC01,

TABLE 3 Linkage disequilibrium (D') among the alleles of the CEH [HLA-B8, SC01, DR3]

HLA-A	HLA-B	HLA-DRB1	HLA-DQA1	HLA-DQB1	D'	p	No. points
1				0201	0.282	<0.05	2
1			0501		0.488	<0.05	2
1		0301			0.505	<0.05	2
1	8				0.706	<0.01	2
	8	0301			0.701	<0.01	2
	8		0501		0.851	<0.01	2
	8			0201	0.935	<0.01	2
		0301	0501		1.000	<0.01	2
			0501	0201	0.554	<0.01	2
1	8	0301			0.490	<0.01	3
1	8		0501		0.509	<0.01	3
	8	0301	0501		0.611	<0.01	3
	8	0301		0201	0.564	<0.01	3
1	8	0301	0501		0.116	n.s.	4
1	8	0301		0201	0.158	n.s.	4
1	8		0501	0201	0.188	n.s.	4
	8	0301	0501	0201	0.040	n.s.	4
1	8	0301	0501	0201	0.058	n.s.	5

DR3] CEH in this dataset was 0.197, of which 70% carried HLA-A1. There were no other instances of HLA-B8 or the complotype SC01. There was a single example of (HLA-DRB1*0301, DQA1*0501, DQB1*0201) without the other markers of the complete CEH and 19% of additional DQA1*0501 (almost all with DRB1*11) and 7% of additional DQB1*0201 (all with DRB1*0701), reflecting the fact that these markers are components of other small blocks. As demonstrated in Table 3, when D' was calculated for more than two markers, D' was decreased, and when calculated for more than three markers, significance was lost. The reason for this is the mathematical necessity to correct for lower-order (HLA-A/B, HLA-B/DR, HLA-DR/DQ, *etc.*) LD. This made sense when the three-point analysis was first applied to the MHC [77] and there was no requirement that the marker alleles even be on the same chromosome. If, however, fixity on one chromosome is the basis of LD in the whole subject (clearly the case), the necessary correction will remove the relevant supporting evidence. These considerations partly explain the distribution of SNP haplotype-defined blocks and the failure of this approach to “see” CEHs. The validity of multipoint LD calculations has been challenged before, and the problem appears to be “general to any kind of analysis of multi-dimensional interactions” [78]. A solution that avoids but does not completely correct this problem is now commonly used and involves sequential (“sliding”) pairwise LD analysis.

Other problems with LD analysis have been described, including the dependence of the LD statistic on the degree of polymorphism of the tested markers [23, 25, 79]. We cannot address all LD technical details here, but one point is clear to us. It is unlikely that these problems

with LD analysis can be overcome simply by improving statistical methodology. There have been a number of cautionary publications about the use of SNP/LD analysis [4, 80–82], and others have suggested technical solutions to some general problems [83]. Furthermore, the null hypothesis for LD analysis is that there is random meiotic crossing over between small blocks and that there is a direct and simple linear relationship between the distance between two genetic loci and the extent of LD. Figure 2 reveals clearly that this null hypothesis is false for the MHC. Others have shown this theoretically [84]. Although not stressed in their publications, the null hypothesis has also been shown to be false by SNP/LD analysis [1, 2].

Even if it were capable of better recognizing large stretches of fixity, SNP/LD analysis cannot in itself reveal the details of haplotype composition, including the nature of CEHs, nor the sizes of fragments of CEHs. Therefore, in the absence of intact CEH/CEH fragment analysis [30], SNP/LD analysis is unlikely to be a great aid in complex disease susceptibility gene localization. Direct haplotype determination within a small genomic region such as the MHC allows one to detect fixity of stretches of DNA of any size directly and unequivocally and independently of assumptions. LD analysis cannot make this important distinction. In fact, at least for the MHC, it is this fixity of DNA that is responsible for significant observed LD, even when it occurs in fragmented blocks.

The MHC haplotype sequencing project [85] is a step in the right direction. Eight consanguineous MHC haplotype homozygotes were chosen for sequencing with complete sequences for HLA-A3, -B7, -DR2 and HLA-A1, -B8, -DR3 already reported [86]. The list more or

less corresponds to that of common CEHs reported by us [10]. Once the sequencing of the present single instances of each haplotype is complete, the project faces the far more daunting task of resequencing multiple independent examples of each CEH to define microvariation on each CEH as well as population representation. These goals are not currently very practical.

One could argue that cost and the difficulties of family studies are prohibitive compared with studies of random individuals. In fact, the limited information from LD analysis of two unrelated subjects yields four limited and guessed haplotypes, whereas the study of three individuals in a family (mother, father, child) yields, in general, four definitive haplotypes, no matter whether SNPs, allotypes, specificities, or alleles are used to define the haplotypes. An advantage of protein allotypes, serologically defined specificities, and nucleotide sequence-defined alleles using cloned DNA over statistically determined SNP haplotypes is that for all of the former, phase is automatically assigned, whereas SNP haplotypes are statistically estimated. At best, well-designed SNP/LD studies of unrelated individuals approach the accuracy of pedigree-determined haplotype studies only when conducted within regions of “low haplotype complexity” similar to HLA-Cw/B, complo-type, or the HLA-DR/DQ block [87]. Higher-order fixity is usually beyond the capability of the methodology.

We believe that the problem with SNP/LD analysis is the statistical nature of the approach rather than the use of SNPs. Although most SNPs tend to be less informative than many recognized MHC allelic differences in terms of allele frequencies and the number of common alleles, they serve potentially useful purposes. For example, it would be of interest to know the extent of variation of SNPs on individual examples of the same CEH. We would predict that such variation is minimal. It would also be of interest to know whether any SNPs or SNP haplotypes are CEH- or CEH variant-specific or common to many CEHs.

There are approaches to disease susceptibility gene localization suggested by and consistent with the existence of CEHs and their fragments. As applied to the MHC (a possible model for other less well-studied parts of the genome), if the MHC susceptibility gene is recessive, the existence of a protective CEH in a patient must be explained. Its presence in a patient could represent a mutation or gene conversion at the susceptibility locus or a historical crossover between the elements of the CEH and the susceptibility locus. The latter would constitute localization information [88].

The haplotype distributions in Figure 2 and the conclusions that can be drawn from them do not require any assumptions, including whether or not the examined populations are in genetic equilibrium (no major modern

population is in genetic equilibrium). It is these directly determined, highly varied MHC haplotypes of varying length, rather than the relatively unvaried short SNP/LD haplotypes common to all populations and guessed at by maximum likelihood statistical methods, that are the genomic reality. It is the observed haplotypes that must be the basis of any rational approach to gene localization.

ACKNOWLEDGMENTS

This work was supported by Grant HL-29583 from the National Heart, Lung, and Blood Institute of the National Institutes of Health. We are indebted to many colleagues for stimulating discussions over many years and for past collaborative studies. We thank George Eisenbarth, Stuart Brink, Kenneth Gabbay, A. Razzaque Ahmed, and Donald Raum in particular.

REFERENCES

1. Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES: High-resolution haplotype structure in the human genome. *Nat Genet* 29:229, 2001.
2. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumensteil B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D: The structure of haplotype blocks in the human genome. *Science* 296:2225, 2002.
3. Wall JD, Pritchard JK: Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet* 4:587, 2003.
4. Kruglyak L, Nickerson DA: Variation is the spice of life. *Nat Genet* 27:234, 2001.
5. Dawson E, Abecasis GR, Bumpstead S, Chen Y, Hunt S, Beare DM, Pabial J, Dibling T, Tinsley E, Kirby S, Carter D, Papaspyridonos M, Livingstone S, Ganske R, Lohmussaar E, Zernant J, Tonisson N, Remm M, Magi R, Puurand T, Vilo J, Kurg A, Rice K, Deloukas P, Mott R, Metspalu A, Bentley DR, Cardon LR, Dunham I: A first-generation linkage disequilibrium map of human chromosome 22. *Nature* 418:544, 2002.
6. Ceppellini R, Siniscalco M, Smith CA: The estimation of gene frequencies in a random-mating population. *Ann Hum Genet* 20:97, 1955.
7. Smith CAB: Counting methods in genetical statistics. *Ann Hum Genet* 21:254, 1957.
8. Raum D, Awdeh Z, Yunis EJ, Alper CA, Gabbay KH: Extended major histocompatibility complex haplotypes in type I diabetes mellitus. *J Clin Invest* 74:449, 1984.
9. Falk CT, Rubinstein P: Haplotype relative risks: an easy way to construct a proper control sample for risk calculations. *Ann Hum Genet* 51:227, 1987.
10. Awdeh ZL, Raum D, Yunis EJ, Alper CA: Extended HLA/complement allele haplotypes: evidence for T/t-like complex in man. *Proc Natl Acad Sci USA* 80:259, 1983.

11. Ahmad T, Neville M, Marshall SE, Armuzzi A, Mulcahy-Hawes K, Crawshaw J, Sato H, Ling K-L, Barnado M, Goldthorpe S, Walton R, Bunce M, Jewell DP, Welsh KI: Haplotype-specific linkage disequilibrium patterns define the genetic topography of the human MHC. *Hum Mol Genet* 12:647, 2003.
12. Fleischnick E, Awdeh ZL, Raum D, Granados J, Alosco SM, Crigler JF Jr, Gerald PS, Giles CM, Yunis EJ, Alper CA: Extended MHC haplotypes in 21-hydroxylase deficiency congenital adrenal hyperplasia: shared genotypes in unrelated patients. *Lancet* 1:152, 1983.
13. Alper CA, Fleischnick E, Awdeh Z, Katz AJ, Yunis EJ: Extended major histocompatibility complex haplotypes in patients with gluten-sensitive enteropathy. *J Clin Invest* 79:251, 1987.
14. Ahmed AR, Yunis EJ, Khatri K, Wagner R, Notani G, Awdeh Z, Alper CA: Major histocompatibility complex haplotype studies in Ashkenazi Jewish patients with pemphigus vulgaris. *Proc Natl Acad Sci USA* 87:7658, 1990.
15. Alper CA, Kruskall MS, Marcus-Bagley D, Craven DE, Katz AJ, Brink SJ, Dienstag JL, Awdeh Z, Yunis EJ: Genetic prediction of nonresponse to hepatitis B vaccine. *N Engl J Med* 321:708, 1989.
16. Hauser SL, Fleischnick E, Weiner HL, Marcus D, Awdeh Z, Yunis EJ, Alper CA: Extended major histocompatibility complex haplotypes in patients with multiple sclerosis. *Neurology* 39:275, 1989.
17. Ahmed AR, Wagner R, Khatri K, Notani G, Awdeh Z, Alper CA, Yunis EJ: Major histocompatibility complex haplotypes and class II genes in non-Jewish patients with pemphigus vulgaris. *Proc Natl Acad Sci USA* 88:5056, 1991.
18. Blumenthal M, Marcus-Bagley D, Awdeh Z, Johnson B, Yunis EJ, Alper CA: HLA-DR2, [HLA-B7, SC31, DR2], and [HLA-B8, SC01, DR3] haplotypes distinguish subjects with asthma from those with rhinitis only in ragweed pollen allergy. *J Immunol* 148:411, 1992.
19. Ahmed AR, Katri K, Yunis EJ, Notani G, Yunis JJ, Awdeh Z, Alper CA: Major histocompatibility complex susceptibility genes for dermatitis herpetiformis compared with those for gluten-sensitive enteropathy. *J Exp Med* 78:2067, 1993.
20. Alper CA, Xu J, Cosmopoulos K, Dolinski B, Stein R, Uko G, Larsen CE, Dubey DP, Densen P, Truedsson L, Sturfelt G, Sjöholm AG: Immunoglobulin deficiencies and susceptibility to infection among homozygotes and heterozygotes for C2 deficiency. *J Clin Immunol* 23:297, 2003.
21. Rubinstein P, Walker M, Carpenter C, Carrier C, Krassner J, Falk C, Ginsberg F: The use of the haplotype relative risk (HRR) and the "haplo-delta" (Dh) estimates in juvenile diabetes from three racial groups [abstract]. *Hum Immunol* 3:384, 1981.
22. Hill WG: Estimation of linkage disequilibrium in randomly mating populations. *Heredity* 33:229, 1974.
23. Lewontin RC: On measures of gametic disequilibrium. *Genetics* 120:849, 1988.
24. Weir BS: *Genetic Data Analysis. Methods for Discrete Population Genetics*. Sunderland, MA, Sinauer Associates, 1990.
25. Lewontin RC: The interaction of selection and linkage. I. General considerations: heterotic models. *Genetics* 49:49, 1964.
26. Hopkins KA: The basic lymphocyte microcytotoxicity tests: standard and AHG enhancement. In Hahn AB, Land GA, Strothman RM (eds): *Laboratory Manual, 4th ed, vol 1*. New York, American Society for Histocompatibility and Immunogenetics, 2000.
27. Reference Protocol: 11th International Histocompatibility Workshop, Yokohama, Japan, 1991.
28. Zetterquist H, Olerup O: Identification of the HLA-DRB1*04, -DRB1*07, and -DRB1*09 alleles by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours. *Hum Immunol* 34:64, 1992.
29. Bunce M, O'Neil C, Barnardo M, Morris P, Welsh K: Phototyping: comprehensive DNA typing for HLA-A, B, C, DRB3, DRB4, DRB5, and DQB1 by PCR with 144 primer mixes utilizing sequence specific primers (PCR-SSP). *Tissue Antigens* 46:355, 1995.
30. Yunis EJ, Larsen CE, Fernandez-Viña M, Awdeh ZL, Romero T, Hansen JA, Alper CA: Inheritable variable sizes of DNA stretches in the human MHC: conserved extended haplotypes and their fragments or blocks. *Tissue Antigens* 62:1, 2003.
31. Marcus-Bagley D, Alper CA: Methods for allotyping complement proteins. In Rose NR, Conway de Macario E, Fahey JL, Friedman H, Penn GM (eds): *Manual of Clinical Laboratory Immunology, 4th ed*. Washington, DC, American Society for Microbiology, 1992.
32. Alper CA, Raum D, Karp S, Awdeh ZL, Yunis EJ: Serum complement 'supergenes' of the major histocompatibility complex in man (complotypes). *Vox Sang* 45:62, 1983.
33. Simon S, Truedsson L, Marcus-Bagley D, Awdeh Z, Eisenbarth GS, Brink SJ, Yunis EJ, Alper CA: Relationship between protein complotypes and DNA variant haplotypes: complotype-RFLP constellations (CRC). *Hum Immunol* 57:27, 1997.
34. Sanchez-Mazas A, Djoulah S, Busson M, Le Monnier de Gouville I, Poirier J-C, Dehay C, Charron D, Excoffier L, Schneider S, Langaney A, Dausset J, Hors J: A linkage disequilibrium map of the MHC region based on the analysis of 14 loci haplotypes in 50 French families. *Eur J Hum Genet* 8:33, 2000.
35. Degli-Esposti MA, Leaver AL, Christiansen FT, Witt CS, Abraham LJ, Dawkins RL: Ancestral haplotypes: conserved population MHC haplotypes. *Hum Immunol* 34:242, 1992.
36. Yunis EJ, Zuñiga J, Larsen CE, Fernández-Viña M, Granados J, Awdeh ZL, Alper CA: Single nucleotide polymorphism blocks and haplotypes: human MHC block

- diversity. In Meyers RA (ed): *Encyclopedia of Molecular Cell Biology and Molecular Medicine*, 2nd ed, vol 13. Weinheim, Wiley-VCH Verlag, 2005.
37. Zhang WJ, Degli-Esposti MA, Cobain TJ, Cameron PU, Christiansen FT, Dawkins RL: Differences in gene copy number carried by different MHC ancestral haplotypes: quantitation after physical separation of haplotypes by pulsed field gel electrophoresis. *J Exp Med* 171:2101, 1990.
 38. Egea GE, Yunis I, Spies T, Strominger T, Awdeh ZL, Alper CA, Yunis EJ: Association of polymorphisms in the HLA-B region with extended haplotypes. *Immunogenetics* 33:4, 1991.
 39. Jongeneel CV, Briant L, Udalova IA, Sevin A, Nedospasov SA, Cambon-Thomsen A: Extensive genetic polymorphism in the human tumor necrosis factor region and relation to extended HLA haplotypes. *Proc Natl Acad Sci USA* 88:9717, 1991.
 40. Singal DP, Blajchman MA: Histocompatibility (HL-A) antigens, lymphocytotoxic antibodies and tissue antibodies in patients with diabetes mellitus. *Diabetes* 22:429, 1973.
 41. Svejgaard A, Platz P, Ryder LP: In Terasaki PI (eds): *Histocompatibility Testing 1980*. Los Angeles, UCLA Tissue Typing Laboratory, 1981.
 42. Solow H, Hidalgo R, Singal DP: Juvenile-onset diabetes: HLA-A, -B, -C, and -DR alloantigens. *Diabetes* 28:1, 1979.
 43. Sachs JA, Cudworth AG, Jaraquemada D, Gorsuch AN, Festenstein H: Type 1 diabetes and the HLA-D locus. *Diabetologia* 18:41, 1980.
 44. McCluskey J, Kay PH, Dawkins RL, Komori K, Christiansen FT, McKann VJ: Association of specific MHC supratypes with rheumatoid arthritis and insulin-dependent diabetes mellitus. *Dis Markers* 1:197, 1983.
 45. Degli-Esposti MA, Abraham LJ, McCann V, Spies T, Christiansen FT, Dawkins RL: Ancestral haplotypes reveal the role of the central MHC in the immunogenetics of IDDM. *Immunogenetics* 36:345, 1992.
 46. Abraham LJ, Leelayuwat C, Grimsley G, Degli-Esposti MA, Mann A, Zhang WJ, Christiansen FT, Dawkins RL: Sequence differences between HLA-B and TNF distinguish different MHC ancestral haplotypes. *Tissue Antigens* 39:117, 1992.
 47. Cascino I, D'Alfonso S, Cappello N, Giordano M, Pugliese A, Awdeh Z, Alper CA, Richiardi PM: Gametic association of HSP70-1 promoter region alleles and their inclusion in extended HLA haplotypes. *Tissue Antigens* 42:62, 1993.
 48. Garcia-Merino A, Alper CA, Usuku K, Marcus-Bagley D, Lincoln R, Awdeh Z, Yunis EJ, Eisenbarth GS, Brink SJ, Hauser SL: Tumor necrosis factor (TNF) microsatellite haplotypes in relation to extended haplotypes, susceptibility to diseases associated with major histocompatibility complex and TNF secretion. *Hum Immunol* 50:11, 1996.
 49. Clavijo OP, Delgado JC, Awdeh ZL, Fici D, Turbay D, Alper CA, Truedsson L, Yunis EJ: HLA-Cw alleles associated with HLA extended haplotypes and C2 deficiency. *Tissue Antigens* 52:282, 1998.
 50. Uglialoro AM, Turbay D, Pesavento PA, Delgado JC, McKenzie FE, Gribben JG, Hartl D, Yunis EJ, Goldfeld AE: Identification of three new single nucleotide polymorphisms in the human tumor necrosis factor- α gene promoter. *Tissue Antigens* 52:359, 1998.
 51. De la Concha EG, Fernandez-Arquero M, Gual L, Vigil P, Martinez A, Urcelay E, Ferreira A, Garcia-Rodriguez MC, Fontan G: MHC susceptibility genes to IgA deficiency are located in different regions on different HLA haplotypes. *J Immunol* 169:4637, 2002.
 52. Pinto C, Smith AG, Larsen CE, Fernández-Viña M, Husain Z, Clavijo OP, Wang ZC, Nisperos B, Hansen JA, Alper CA, Yunis EJ: HLA-Cw*0409N is associated with HLA-A*2301 and HLA-B*4403-carrying haplotypes. *Hum Immunol* 65:181, 2004.
 53. Baena A, Leung JY, Sullivan AD, Landires I, Vasquez-Luna N, Quiñones-Berrocal J, Fraser PA, Uko GP, Delgado JC, Clavijo OP, Thim S, Meshnick SR, Nyirenda T, Yunis EJ, Goldfeld AE: TNF- α promoter single nucleotide polymorphisms are markers of human ancestry. *Genes Immun* 3:482, 2002.
 54. Shao W, Lobashevsky ES, Kaslow RA, Dorak MT: MICA intron 1 sequences of conserved extended HLA haplotypes: implications for sequencing-based typing. *Genes Immun* 5:371, 2004.
 55. Awdeh ZL, Alper CA, Eynon E, Alosco SM, Stein R, Yunis EJ: Unrelated individuals matched for MHC extended haplotypes and HLA-identical siblings show comparable responses in mixed lymphocyte culture. *Lancet* 2:853, 1985.
 56. Awdeh ZL, Alper CA, Fici DA, Ronco II P, Yunis EJ: Predictability of alloreactivity among unrelated individuals: role for HLA-DPB1. *Tissue Antigens* 46:180, 1995.
 57. Alper CA, Awdeh Z: Incomplete penetrance of MHC susceptibility genes: prospective analysis of polygenic traits. *Tissue Antigens* 56:199, 2000.
 58. Alper CA, Marcus-Bagley D, Awdeh Z, Kruskall MS, Eisenbarth GS, Brink SJ, Katz AJ, Stein R, Bing DH, Yunis EJ, Schur PH: Prospective analysis suggests susceptibility genes for deficiencies of IgA and several other immunoglobulins on the [HLA-B8, SC01, DR3] conserved extended haplotype. *Tissue Antigens* 56:207, 2000.
 59. Calvo B, Castaño L, Marcus-Bagley D, Fici DA, Awdeh Z, Alper CA: The [HLA-B18, F1C30, DR3] conserved extended haplotype carries a susceptibility gene for IgD deficiency. *J Clin Immunol* 20:216, 2000.
 60. Matsui Y, Alosco SM, Awdeh Z, Duquesnoy RJ, Page PL, Hartzman RJ, Alper CA, Yunis EJ: Linkage disequilibrium of *HLA-SB1* with the *HLA-A1, B8, DR3, SC01* and of *HLA-SB4* with the *HLA-A26, Bw38, Dw10, DR4, SC21* extended haplotypes. *Immunogenetics* 20:623, 1984.

61. Yunis JJ, Salazar M, Delgado MB, Alper CA, Bing DH, Yunis EJ: HLA-DQA1, DQB1 and DPB1 alleles on HLA-DQ2- and DQ9-carrying extended haplotypes. *Tissue Antigens* 41:37, 1993.
62. Welch TR, Beischel L, Balakrishnan K, Quinlan M, West CD: Major-histocompatibility-complex extended haplotypes in membranoproliferative glomerulonephritis. *N Engl J Med* 314:1476, 1986.
63. Patrick A, Briggs D, Welsh K, Jacoby R, Williamson E, Jones V: Extended haplotypes in rheumatoid arthritis and preliminary evidence for an interaction with immunoglobulin genes. *Dis Markers* 4:139, 1986.
64. Singal DP, Green D, D'Souza M, Reid B: Polymorphism of three HLA-DR7 bearing major histocompatibility complex extended haplotypes. *Mol Immunol* 26:705, 1989.
65. Degli-Esposti MA, Andreas A, Christiansen FT, Schalke B, Albert E, Dawkins RL: An approach to the localization of the susceptibility gene for generalized myasthenia gravis by mapping recombinant ancestral haplotypes. *Immunogenetics* 35:355, 1992.
66. Larsen CE, Alper CA: The genetics of HLA-associated disease. *Curr Opin Immunol* 16:660, 2004.
67. Volanakis JE, Zhu ZB, Schaffer FM, Macon KJ, Palermos J, Barger BO, Go R, Campbell RD, Schroeder HW Jr, Cooper MD: Major histocompatibility complex class III genes and susceptibility to immunoglobulin A deficiency and common variable immunodeficiency. *J Clin Invest* 89:1914, 1992.
68. Schroeder HW Jr, Zhu ZB, March RE, Campbell RD, Berney SM, Nedospasov SA, Turetskaya RL, Atkinson TP, Go RC, Cooper MD, Volanakis JE: Susceptibility locus for IgA deficiency and common variable immunodeficiency in the HLA-DR3, -B8, -A1 haplotypes. *Mol Med* 4:72, 1998.
69. Vorechovsky I, Cullen M, Carrington M, Hammarström L, Webster ADB: Fine mapping of *IGAD1* in IgA deficiency and common variable immunodeficiency: identification and characterization of haplotypes shared by affected members of 101 multiple-case families. *J Immunol* 164:4408, 2000.
70. Walsh EC, Mather KA, Schaffner SF, Farwell L, Daly MJ, Patterson N, Cullen M, Carrington M, Bugawan TL, Erlich H, Campbell J, Barrett J, Miller K, Thomson G, Lander ES, Rioux JD: An integrated haplotype map of the human major histocompatibility complex. *Am J Hum Genet* 73:580, 2003.
71. Jeffreys AJ, Kauppi L, Neumann R: Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nat Genet* 29:217, 2001.
72. Cullen M, Perfetto SP, Klitz W, Nelson G, Carrington M: High resolution patterns of meiotic recombination across the human major histocompatibility complex. *Am J Hum Genet* 71:759, 2002.
73. Thompson EA, Neel JV: Allelic disequilibrium and allele frequency distribution as a function of social and demographic history. *Am J Hum Genet* 60:197, 1997.
74. Bodmer WF, Bodmer JG: Evolution and function of the HLA system. *Br Med Bull* 34:309, 1978.
75. Lewontin RC: *The Genetic Basis of Evolutionary Change*. New York, Columbia Univ. Press, 1974.
76. Bodmer JG, Kennedy LJ, Lindsay J, Wasik AM: Applications of serology and the ethnic distribution of three locus HLA haplotypes. *Br Med Bull* 43:94, 1987.
77. Piazza A: Haplotypes and Linkage Disequilibria from Three-locus Phenotypes. In: Kissmeyer-Nielsen F (ed): *Histocompatibility Testing*, 1975. Copenhagen, Munksgaard, 1975.
78. Baur MP: Haplotype Frequencies and Linkage Disequilibria. In Ohayon E, Cambon-Thomsen A (eds): *Colloque INSERM Génétiques des Population Humaines*. Paris, Les Editions INSERM, 1985.
79. Zondervan KT, Cardon LR: The complex interplay among factors that influence allelic association. *Nat Rev Genet* 5:89, 2004.
80. Terwilliger JD, Weiss KM: Linkage disequilibrium mapping of complex disease: fantasy or reality? *Curr Opin Biotechnol* 9:578, 1998.
81. Terwilliger JD, Haghghi F, Hiekkalinna TS, Göring HHH: A *bias*-ed assessment of the use of SNPs in human complex traits. *Curr Opin Genet Dev* 12:726, 2002.
82. Spence MA, Greenberg DA, Hodge SE, Vieland VJ: The emperor's new methods. *Am J Hum Genet* 72:1084, 2003.
83. Newton-Cheh C, Hirschhorn JN: Genetic association studies of complex traits: design and analysis issues. *Mutat Res* 573:54, 2005.
84. Nordborg M, Tavaré S: Linkage disequilibrium: what history has to tell us. *Trends Genet* 18:83, 2002.
85. Allcock RJN, Atrazhev AM, Beck S, de Jong PJ, Elliott JF, Forbes S, Halls K, Horton R, Osoegawa K, Rogers J, Sawcer S, Todd JA, Trowsdale J, Wang Y, Williams S: The MHC haplotyping project: a resource for HLA-linked association studies. *Tissue Antigens* 59:520, 2002.
86. Stewart CA, Horton R, Allcock RJ, Ashurst JL, Atrazhev AM, Coghill P, Dunham I, Forbes S, Halls K, Howson JM, Humphray SJ, Hunt S, Mungall AJ, Osoegawa K, Palmer S, Roberts AN, Rogers J, Sims S, Wang Y, Wilming LG, Elliott JF, de Jong PJ, Sawcer S, Todd JA, Trowsdale J, Beck S: Complete MHC haplotype sequencing for common disease gene mapping. *Genome Res* 14:1176, 2004.
87. Kirk KM, Cardon LR: The impact of genotyping error on haplotype reconstruction and frequency estimation. *Eur J Hum Genet* 10:616, 2002.
88. Erlich H, Griffith RL, Bugawan TL, Ziegler R, Alper C, Eisenbarth G: Implication of specific DQB1 alleles in genetic susceptibility and resistance by identification of IDDM siblings with novel HLA-DQB1 allele and unusual DR2 and DR1 haplotypes. *Diabetes* 40:478, 1991.