Am. J. Hum. Genet. 71:585–594, 2002

Dissecting the Genetic Complexity of the Association between Human Leukocyte Antigens and Rheumatoid Arthritis

Damini Jawaheer,[1] Wentian Li,[1] Robert R. Graham,[2] Wei Chen,[3] Aarti Damle,[1] Xiangli Xiao,[1] Joanita Monteiro,[1] Houman Khalili,[1] Annette Lee,[1] Robert Lundsten,[1] Ann Begovich,[4] Teodorica Bugawan,[4] Henry Erlich,[4] James T. Elder,[5,6] Lindsey A. Criswell,[7] Michael F. Seldin,[8] Christopher I. Amos,[3] Timothy W. Behrens,[2] and Peter K. Gregersen[1]

[1]Center for Genomics and Human Genetics, North Shore–Long Island Jewish Research Institute, Manhasset, New York; [2]Center for Immunology, University of Minnesota, Minneapolis; [3]Departments of Epidemiology and Biomathematics, University of Texas, M. D. Anderson Cancer Center, Houston; [4]Roche Molecular Systems, Berkeley, CA; Departments of [5]Dermatology and [6]Radiation Oncology, University of Michigan, Ann Arbor; [7]The Rosalind Russell Medical Research Center for Arthritis, Department of Medicine, Division of Rheumatology, University of California at San Francisco, San Francisco; and [8]Department of Biological Chemistry, University of California at Davis, Davis

Rheumatoid arthritis (RA) is an inflammatory disease with a complex genetic component. An association between RA and the human leukocyte antigen (HLA) complex has long been observed in many different populations, and most studies have focused on a direct role for the HLA-DRB1 "shared epitope" in disease susceptibility. We have performed an extensive haplotype analysis, using 54 markers distributed across the entire HLA complex, in a set of 469 multicase families with RA. The results show that, in addition to associations with the DRB1 alleles, at least two additional genetic effects are present within the major histocompatibility complex. One of these lies within a 497-kb region in the central portion of the HLA complex, an interval that excludes DRB1. This genetic risk factor is present on a segment of a highly conserved ancestral A1-B8-DRB1*03 (8.1) haplotype. Additional risk genes may also be present in the HLA class I region in a subset of DRB1*0404 haplotypes. These data emphasize the importance of defining haplotypes when trying to understand the HLA associations with disease, and they clearly demonstrate that such associations with RA are complex and cannot be completely explained by the DRB1 locus.

Introduction

Rheumatoid arthritis (RA [MIM 180300]) is an inflammatory disease with autoimmune features, the etiology of which is still unknown. An association between RA and human leukocyte antigens (HLA) was first demonstrated in the 1970s, when the mixed lymphocyte culture (MLC) type Dw4 was found to be present at increased frequencies among patients with RA compared with control individuals (Stastny 1978). Since then, a large number of population studies have confirmed the association between RA and HLA-DRB1*04 (Ollier and Thomson 1992). The size of the genetic component in RA has been computed using the relative recurrence risk for siblings of probands with RA (λ_s); it is likely that this value lies between 5 and 10 (Seldin et al. 1999; Jawaheer et al. 2001). It has been estimated that approximately one-third to one-half of this total genetic contribution in RA can be attributed to genes in the HLA complex, also known as the "major histocompatibility complex" (MHC) (Deighton et al. 1989). In two separate genomewide screens using affected sibling pairs with RA, it was demonstrated that HLA has the largest genetic contribution in RA, and the relative contribution of HLA genes (λ_{HLA}) was found to be 1.7–1.8 (Cornelis et al. 1998; Jawaheer et al. 2001). However, despite >20 years of research to unravel the role of HLA in RA, the exact cause of this association remains unknown.

The vast majority of studies have focused on a direct role for DRB1 alleles that encode a common structural element, designated the "shared epitope" (Gregersen et al. 1987). However, it is unlikely that the shared epitope alone can completely explain the HLA-associated risk for RA (Moxley and Cohen 2002). Different shared epitope-encoding DRB1 alleles vary in the strength of their association with RA, suggesting additional allelic or haplotypic effects on risk. In addition, particular heterozygote combinations that have identical combinations of shared epitope sequences may, nevertheless, differ dramatically in their relative risk (RR) for RA. For example, a RR of 6.4 is associated with the DRB1*0101/

Received April 23, 2002; accepted for publication June 14, 2002; electronically published August 9, 2002.

Address for correspondence and reprints: Dr. Peter K. Gregersen, Center for Genomics and Human Genetics, North Shore-LIJ Research Institute, 350 Community Drive, Manhasset, NY 11030. E-mail: peterg@nshs.edu

0002-9297/2002/7103-0013$15.00

Table 1

Structure of the Multiplex Families with RA

No. of Parents Available	No. of Families (n = 469)	No. of Families with			
		2 Affected Sibs	3 Affected Sibs	4 Affected Sibs	6 Affected Sibs
2	59	56	3	0	0
1	164	146	18	0	0
0	246	212[a]	29[b]	4[c]	1[d]

[a] Two families had one unaffected sibling genotyped.
[b] Three families had one, two, and five unaffected siblings genotyped.
[c] One family had five unaffected siblings genotyped.
[d] One family had one unaffected sibling genotyped.

0401 genotype, but, for the DRB1*0401/0404 genotype, the RR is 31.3 (Hall et al. 1996). The HLA shared epitope does not show an association with RA among African-American and Hispanic-American RA populations (McDaniel et al. 1995; Teller et al. 1996). Moreover, when the marker association-segregation χ^2 method is used to investigate the HLA component in RA, a role for the shared epitope in RA is rejected (Dizier et al. 1993).

The HLA complex occupies a 3.6-Mb region on the short arm of chromosome 6 (6p21.3 band) and contains ~220 defined genes (Milner and Campbell 2001). Many of the genes in this region are involved in immune function. However, aside from DRB1, the only other locus within the MHC that has been reported as a possible candidate for RA susceptibility is the tumor necrosis factor (TNF) locus (Mulcahy et al. 1996; Ota et al. 2001; Waldron-Lynch et al. 2001). In the present study, we have used 54 polymorphic markers to examine HLA haplotypes in 469 multiplex families with RA (Gregersen 1998; Jawaheer et al. 2001). This has allowed us to define additional genetic regions within the MHC, in addition to DRB1, that confer risk of RA. These data emphasize the importance of defining haplotypes when trying to understand the HLA associations with disease, and they clearly demonstrate that these associations with RA are complex and cannot be completely explained by the DRB1 locus.

Subjects and Methods

Study Populations

Participants in the present study were members of 469 white families with multiple cases of RA who were drawn from the North American Rheumatoid Arthritis Consortium (NARAC) collection. The number of parents and affected siblings who participated in the study are shown in table 1. Parental DNA samples were available for typing in 223 families. Among those families for which parental DNA was unavailable, 7 families had DNA from unaffected siblings of the proband. The clinical characteristics of these families have been described elsewhere (Gregersen 1998; Jawaheer et al. 2001).

Laboratory Procedures

DNA samples from the RA multicase families were isolated from the mononuclear peripheral blood cells, using a salting-out kit (BIO-101). Fifty-three microsatellite markers distributed throughout the HLA complex were used (fig. A [online only] and table A [online only]). These included 5 TNF markers (TNFa–e) (Udalova et al. 1993), 35 HLA markers described elsewhere (Nair et al. 2000), and 13 markers (labeled "MN6Sxxxx") that were developed in a collaboration between the Gregersen and Behrens laboratories (Graham et al. 2002 [in this issue]). The exact marker locations were determined from University of California, Santa Cruz, Genome Bioinformatics Web site and the Web site of the Sanger Institute. Fluorescent genotyping was performed using amplification in 96-well plates under standard PCR conditions (as described on the Web site of the Center for Medical Genetics, Marshfield Medical Research Foundation), pooling of PCR products using a Tecan MiniPrep 75, and electrophoresis of the pooled PCR products on a 3700 DNA Analyzer (Applied Biosystems). Three of the markers, TNFb, D6S2669, and D6S2681, required an annealing temperature of 66°C, instead of 55°C, in the PCR.

Broad-level HLA-DRB1 typing and high-resolution DRB1*04 typing were accomplished by initial PCR amplification of groups of alleles (all DRB1 alleles for broad-level typing and group-specific amplification for DRB1*04 alleles) using biotinylated PCR primers, followed by hybridization to immobilized sequence-specific oligonucleotide probes in a linear-array format. Positive hybridization reactions were detected using a streptavidin-horseradish peroxidase conjugate and a soluble colorless substrate, 3,3′, 5,5′-tetramethylbenzidine (Erlich et al. 1991). A computer algorithm, based on the sequence-specific oligonucleotide–probe hybridization pattern, and the A. Nolan Immunogenetics/HLA Sequence Database (1999) were used to assign genotypes to each sample.

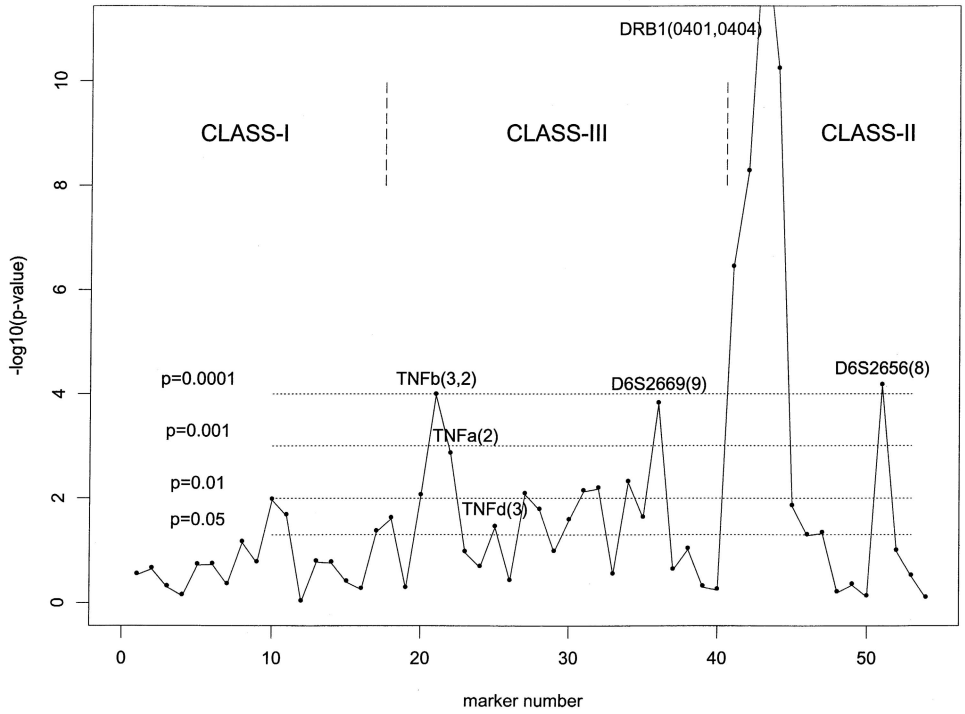


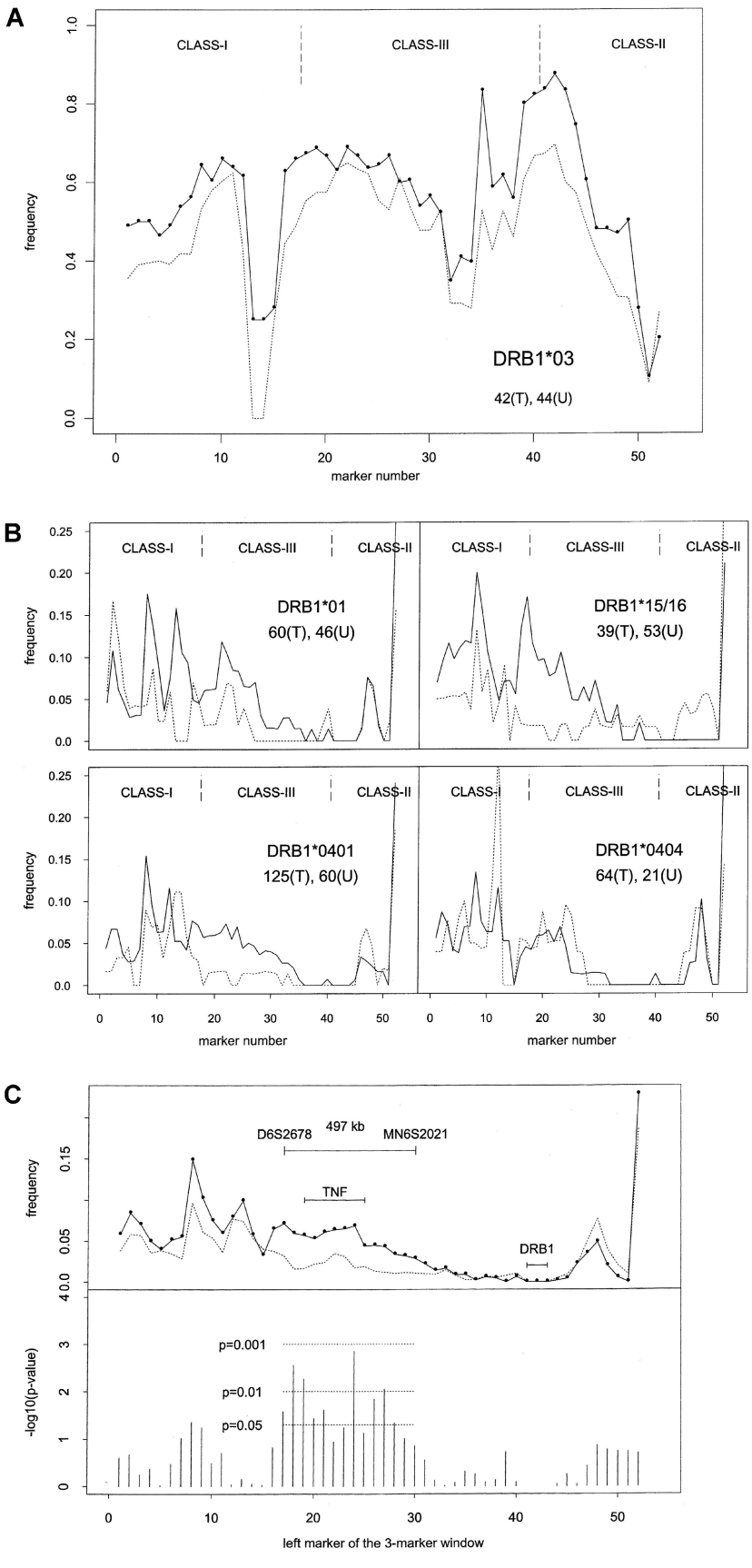
Figure 1 Results of the single-marker PDT analysis. The $-\log_{10}$ of the *P* value obtained from the global PDT (*Y*-axis) is plotted against the order in which the markers are arranged across the HLA complex (*X*-axis). In addition to DRB1, four markers within the class III region show a significant association ($P < .005$) with RA. The association with marker D6S2656 (allele 8), just centromeric to DRB1, reflects strong linkage disequilibrium with DRB1*04 alleles. However, as discussed in the text, the marker associations at TNFb (allele 3), TNFa (allele 2), TNFd (allele 3), and D6S2669 (allele 9) reflect an association with a portion of the ancestral A1-B8-DRB1*03 (8.1) haplotype. See table A to correlate marker numbers with marker names and actual position (fig. A).

Data Analysis

The raw data were analyzed using the GENESCAN analysis (version 3.5) and GENOTYPER (version 2.1) software packages (Applied Biosystems). Each genotype was read independently by two individuals, to ensure accurate calling of alleles. PedCheck was used to identify any inheritance errors in the family data.

Association analysis.—To test for association, we used a modified version of the Pedigree Disequilibrium Test (PDT) (P. Stuart, personal communication), a family-based association analysis program using untransmitted alleles as control samples and transmitted alleles as cases (Martin et al. 2000); it considers parent–affected child triads and discordant sib pairs as the basis unit. The modified version of the PDT program includes parent–affected child dyads as well as triads and discordant sib pairs. In the PDT test result, the overall *P* value, as well as the *P* value for testing uneven frequencies in transmitted and untransmitted groups for an individual allele, are reported.

Haplotype analysis.—We used the GENEHUNTER program to reconstruct haplotypes from the genotyping data from the 469 families with RA (Kruglyak et al. 1996). To ensure the accuracy of these haplotypes, the SIMWALK program (Sobel and Lange 1996) was also used for haplotype reconstruction. A Perl program (W. Li, unpublished data) was then utilized for discovery of haplotype blocks. Detection of conserved haplotype patterns among the haplotypes we reconstructed was achieved by finding the most common three-marker minihaplotype at each window position (1-2-3, 2-3-4,…, and 52-53-54). When such a three-marker minihaplotype can be extended to the neighboring window, it indicates the existence of a longer conserved haplotype. For any particular minihaplotype of interest (e.g., portion of the ancestral A1, B8, DRB1*03 8.1 [8.1] haplotype), their frequencies in two groups (i.e., transmitted and untransmitted haplotypes) can be tested by a simple χ^2 test. When the sample sizes are small, Fisher's exact test can be used to replace the χ^2 test. If the two groups are case and control genotypes, the haplotype information is unknown. In that case, the presence of a minihaplotype of interest is determined by a partial presence of the marker allele in the genotype. This method will tend to overestimate control haplotype frequencies and thus is a conservative approach for identification of positively asso-


A
CLASS-I
CLASS-III
CLASS-II
frequency
DRB1*03
42(T), 44(U)
marker number
B
DRB1*01
60(T), 46(U)
DRB1*15/16
39(T), 53(U)
DRB1*0401
125(T), 60(U)
DRB1*0404
64(T), 21(U)
C
D6S2678
497 kb
MN6S2021
TNF
DRB1
p=0.001
p=0.01
p=0.05
-log10(p-value)
left marker of the 3-marker window

ciated allelic combinations found on minihaplotypes of adjacent markers.

Calculations of entropy.—To characterize the diversity of haplotypes, we calculated an entropy measure of the three-marker minihaplotypes. In this measure, frequency counts of all minihaplotypes were determined $\{p_i > 0\}$ $(i = 1,2,\ldots,m)$, where m is the number of minihaplotypes observed. The normalized entropy is defined as $E = -\sum_i p_i \log(p_i)/\log(m)$ if $m > 1$, and $E = 0$, if $m = 1$, and is a measure of the diversity of minihaplotypes in the "observed data." If all minihaplotypes are evenly distributed, then $E = 1$. On the other hand, if a single haplotype is far more frequent than other haplotypes, then E approaches 0. It can be shown that the normalized entropy is related to the likelihood L of a multinomial distribution as $L = \exp[-N \log(m)E]$, where N is the sample size.

Results

To further investigate which portion(s) of the HLA complex are associated with RA, we used genetic markers distributed across the MHC to examine patterns of transmission disequilibrium in a panel of 469 families with RA that included affected sibling pairs. Each of these markers was independently tested for disease association, using a modified version of the Pedigree Disequilibrium Test (PDT), which included dyads as well as triads in the analysis (see the "Subjects and Methods" section). As expected, the strongest association was observed with DRB1 (DRB1*0401, $P = 7 \times 10^{-9}$; DRB1*0404, $P = 6 \times 10^{-5}$). However, as shown in figure 1, a number of markers in the HLA class III region also showed a significant association with RA. Among these were TNFb (allele 2 [$P = 8.15 \times 10^{-3}$] and allele 3 [$P = 3.33 \times 10^{-3}$]), TNFa (allele 2 [$P = 9.67 \times 10^{-5}$]), TNFd (allele 3 [$P = 1.12 \times 10^{-3}$]), and D6S2669 (allele 9 [$P = 3.28 \times 10^{-4}$].

Simple inspection of the data showed that the combination of TNF b3-a2-d3 and D6S2669 allele 9 formed a common haplotype. We anticipated that this haplotype would be in linkage disequilibrium with one or more HLA-DRB1 alleles, possibly DRB1*0401 or *0404. Using the GENEHUNTER analysis package, we reconstructed full haplotypes wherever possible from the available genotyping data. Surprisingly, after selecting for haplotypes bearing the TNF b3-a2-d3 and D6S2669-9 allele combination among the probands, we found that the DRB1*0301 allele (*not* DRB1*0401 or *0404) was present on 52 (95%) of the 55 haplotypes. When selecting for only those haplotypes with the TNF allele combination TNF b3-a2-d3 (without allele 9 of D6S2669), 66 (61%) of the 108 haplotypes had a DRB1*0301 allele. Furthermore, we observed that the allelic combination of TNF b3-a2-d3 and D6S2669-9 is part of a conserved ancestral DRB1*0301-bearing haplotype (A1-B8-DRB1*03, also known as the 8.1 haplotype), which occurs frequently in white individuals (Price et al. 1999; Graham et al. 2002 [in this issue]).

To determine whether portions of the 8.1 ancestral haplotype are present on haplotypes without DRB1*0301, all haplotypes were grouped according to their broad HLA-DRB1 type. In addition, within each group, the haplotypes were stratified as transmitted or untransmitted, depending on whether they were transmitted to an affected offspring. We then examined a three-marker "window" of allelic combinations by consecutively moving this window across the entire HLA complex (1-2-3, 2-3-4,..., and 52-53-54). For each DRB1 haplotype group, the percent similarity of each three-marker minihaplotype to the consensus 8.1 haplotype was determined. As can be seen from the pattern shown in figure 2*A*, among the DRB1*0301-bearing haplotypes, the transmitted haplotypes shared a portion of the consensus 8.1 haplotype more often than did the untransmitted haplotypes. Furthermore, as shown in figure 2*B*, in several haplotypes that do not bear DRB1*0301, there is still a tendency for the transmitted haplotypes to contain segments of the ancestral 8.1 haplotype, particularly in the midportion of the HLA complex. (The data for all the major DRB1 haplotype groups are shown in figure B [online only]). When all the non-DRB1*0301 haplotypes were combined and analyzed as a group (fig. 2*C*), the three-marker minihaplotypes matching the ancestral al 8.1 haplotype were present significantly more frequently on transmitted haplotypes than on untransmitted haplotypes. Overall, evidence for linkage disequilibrium with an RA disease-susceptibility locus extends from marker D6S2678 to marker MN6S2021 on this haplotype. Notably, this re-

Figure 2 A three-marker "moving window" used to assess the similarity of transmitted (*solid line; with dots* in panels *A* and *C*) and untransmitted (*dashed line*) haplotypes to the ancestral 8.1 haplotype consensus sequence. The frequency of three-marker windows matching the "consensus" 8.1 haplotype sequence (*Y*-axis) is plotted against marker positions (*X*-axis). Marker positions and corresponding marker names are listed in table A. The sample sizes for the transmitted (T) and untransmitted (U) haplotypes are also shown. *A,* When all haplotypes have DRB1*0301, the transmitted haplotypes are more similar to the consensus sequence than are the untransmitted haplotypes. *B,* The same pattern is observed for DRB1*01, *15/*16, and *0401 but not for *0404. *C,* The combined data from all non-DRB1*0301 haplotypes are summarized (*top panel*) (details shown in figure B). It also shows that the transmitted haplotypes are significantly different ($P < .05$) (*bottom panel*) from the untransmitted ones in the interval between D6S2678 and MN6S2021.

gion does not encompass the DRB1 locus, an observation consistent with the fact that the DRB1*0301 allele itself shows no association with RA ($P = .62$ for single-marker PDT). Interestingly, this risk region does not include the D6S2669 marker either, which showed a significant association with RA in the PDT analysis.

We also investigated whether there are any additional factors on the transmitted haplotypes within each DRB1 group that might be different from the untransmitted haplotypes. In order to address this question in a general way, we used a novel approach of measuring the overall diversity of the transmitted and untransmitted haplotypes, in terms of the normalized entropy (E) (Shannon 1948). If combinations of alleles are fully randomized in a collection of haplotypes, $E = 1$. Conversely, if a single haplotype dominates the collection, E approaches 0. In most cases, there was no significant difference in the E value between the transmitted and untransmitted haplotypes (ΔE value), as shown in figure 3 (top panel) for the DRB1*01 haplotypes (also see fig. C [online only]). However, the transmitted DRB1*0404 haplotypes were less diverse than the untransmitted ones (fig. 3, bottom panel), suggesting that a conserved portion of the DRB1*0404 haplotype could harbor a genetic risk factor for RA that is independent of, or interacting with, DRB1*0404.

As shown in figure 2*B*, the transmitted DRB1*0404 haplotypes do not preferentially bear portions of the 8.1 ancestral haplotype. Therefore, we looked for other consensus sequences that might associate with RA on these DRB1*0404 haplotypes. Interestingly, a group of three-marker haplotypes in the class I region are preferentially found on transmitted haplotypes (see table 2). These data indicate that a region of ~659 kb contains a gene that modifies the degree of risk conferred by DRB1*0404 haplotypes. This region (from marker D6S2703 to D6S2696) is located just centromeric to the HLA-A locus. Thus, it does not overlap with the risk region found on the ancestral 8.1 haplotype.

The analyses described above were performed on haplotypes generated by the GENEHUNTER program, using only families in which these haplotypes could be clearly defined. A comparison of the haplotypes generated by GENEHUNTER with those generated by

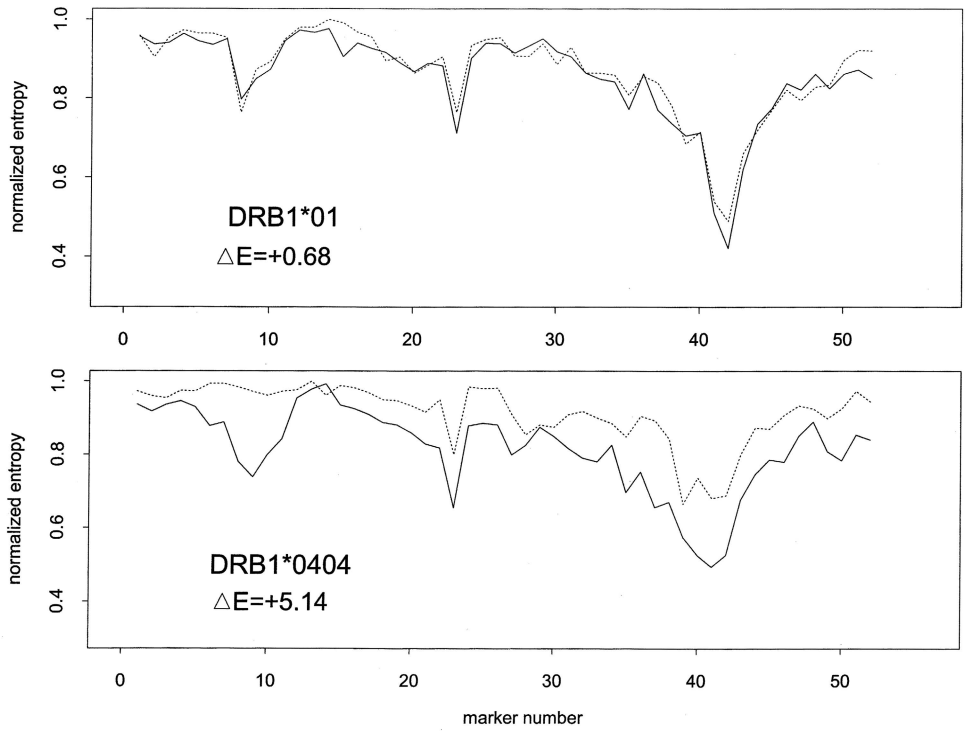


Figure 3 Normalized entropy plots for the transmitted (*solid line*) and untransmitted (*dashed line*) haplotypes in the DRB1*01 and *0404 haplotype groups. (For marker names, see table A.) The ΔE value given for each plot reflects the difference in diversity between untransmitted and transmitted haplotypes. DRB1*01 haplotypes do not display a large difference in diversity between transmitted and untransmitted haplotypes. By contrast, for the DRB1*0404 haplotypes, ΔE is strongly positive, and this reflects the much lower diversity (lower E value) among transmitted haplotypes compared to the untransmitted haplotypes. As discussed in the text and shown in figure 3, the lower diversity among DRB1*0404 transmitted haplotypes cannot be explained by the presence of sequences derived from the 8.1 conserved haplotype. E values for other haplotype groups appear in figure C.

Table 2

Comparison of Three-Marker Minihaplotypes in Transmitted vs. Untransmitted DRB1*0404 Haplotypes in Families with RA

Marker Alleles at[a]							Match/Total in		
M7	M8	M9	M10	M11	M12	M13	Transmitted	Untransmitted	*P*
21	4	5					16/72	0/20	.019
	4	5	2				19/67	0/20	.0048
		5	2	6			29/66	3/23	.011
			2	6	1		20/63	2/22	.048
				6	1	2	17/63	1/24	.019

Note.— Five consecutive three-marker windows (7-8-9, ..., 11-12-13) corresponding to the interval between markers D6S2702 (marker 7) and D6S2691 (marker 13) showed a significant difference between the number of transmitted DRB1*0404 haplotypes that match the DRB1*0404 consensus haplotype and the number of untransmitted haplotypes that match the consensus.

[a] Markers are as follows: M7 = D6S2702, M8 = D6S2700, M9 = D6S2699, M10 = D6S2698, M11 = D6S2697, M12 = D6S2696, and M13 = D6S2691.

SIMWALK showed that, for pedigrees with at least one parent typed, the reconstructed haplotypes from the two programs were very similar. In contrast, for pedigrees with none of the parents typed, the two programs often led to quite different reconstructed haplotypes. Therefore, for minihaplotype analyses, we generally restricted ourselves to those pedigrees with one or two parents typed ($n = 223$; see table 1), with clearly defined haplotypes. The results of the three-marker minihaplotype analyses using haplotypes constructed by GENEHUNTER (figs. 2*A*–2*C* and 3) were highly similar to those generated using SIMWALK haplotypes (data not shown).

Discussion

In the present study, we have comprehensively analyzed the role of MHC haplotypes in genetic susceptibility for RA. To perform these analyses, we were fortunate to have available one of the largest existing collections of multiplex families with RA, assembled over the last five years by the North American Rheumatoid Arthritis Consortium (Gregersen 1998). Overall, the results strongly suggest that considerable genetic heterogeneity underlies the linkage of HLA to this disease and that at least two genetic regions other than HLA-DRB1 contribute to risk of RA.

First, the data show that the central portion of the ancestral 8.1 (A1-B8-DRB1*03) haplotype contains one or more risk factors for RA. This is consistent with previous reports that marker alleles in the TNF region may be associated with RA (Mulcahy et al. 1996; Ota et al. 2001; Waldron-Lynch et al. 2001), since the TNF gene is present within the risk haplotype that we have defined. Several of these studies have suggested that the RA-associated TNF alleles may be found on a DRB1*03 haplotype (Mulcahy et al. 1996; Zanelli et al. 2001). However, to our knowledge, the present analysis is the first study to comprehensively define the extent of this risk haplotype and its relationship to the ancestral 8.1 haplotype. Our data indicate that risk genes could be present anywhere in a relatively large genomic region extending from HLA-C to the C4 locus. Maximal significance appears to be localized around the TNF gene complex in the family-based association analysis reported here. However, a separate case-control analysis (data not shown) continues to show a broad region of association extending to ~500 kb. Nevertheless, it is of interest that high TNF production is associated with TNF alleles found on the ancestral 8.1 haplotype (Wilson et al. 1997), and this may explain the association observed with RA. A recent study also raised the possibility that HLA-C alleles could be risk factors for some forms of RA (Yen et al. 2001). However, there are ~50 genes in this region of interest (MCH Sequencing Consortium 1999), and a significant proportion of these genes are related to immune function. Therefore, it is difficult to determine which genes in this region are actually responsible for the association we have observed.

The 8.1 haplotype is carried by most white individuals with HLA-B8 and has been associated with a variety of interesting immunological diseases and phenotypes (Price et al. 1999). These include common variable immunodeficiency (MIM 240500) and immunoglobulin A (IgA) deficiency (MIM 137100), as well as a number of autoimmune diseases, such as systemic lupus erythematosus (MIM 152700]), insulin-dependent diabetes mellitus (MIM 222100) and autoimmune hepatitis (MIM 142395). In the setting of vaccination or viral infection, the A1-B8-DRB1*03 haplotype has been associated with relative lack of responsiveness (Alper et al. 1989), as well as accelerated progression of HIV disease (Kaslow et al. 1990). Interestingly, this may extend to immune responses to Epstein-Barr–virus infec-

tion (Caruso et al. 1993), which has been proposed as a potential etiologic agent in RA and other autoimmune diseases (Tosato et al. 1981; James et al. 2001). These findings may reflect regulation of the immune response by classical HLA molecules. However, a variety of more nonspecific immunological parameters have also been associated with the 8.1 haplotype (Price et al. 1999), particularly involving the regulation of cytokine production (Candore et al. 1994; Lio et al. 2001). Overall, there appears to be something immunologically distinct about the 8.1 haplotype, and our data suggest that the special immunological properties of the 8.1 haplotype may contribute to the pathogenesis of RA.

Our analysis has also revealed evidence for a second genetic effect on RA susceptibility within the class I region of some DRB1*0404 haplotypes. Subsequent to the identification of DRB1*0404 as a risk factor for RA in the 1980s (Nepom et al. 1986), there have been hints that DRB1*0404 haplotypes may play a distinct role in susceptibility to RA. In particular, although the DRB1*0404 allele alone appears to confer a modest risk for RA (Ronningen et al. 1992; Thomson et al. 1999), the compound heterozygote DRB1*0401/*0404 displays very high RR (Hall et al. 1996). The results of the present study confirm a high RR for the compound heterozygote DRB1*0401/0404 (RR = 19.32; 95% CI = 6–63). In addition, a subset of DRB1*0404 haplotypes contain a common set of minihaplotypes in the class I region, which confers risk independent of the DRB1*0404 allele (table 2). The region of interest extends over 600 kb and contains ~12 functional genes, including HLA-E, a ligand for members of the immunoglobulin-like killer inhibitory receptor family, as well as other candidates of potential interest (MHC Sequencing Consortium 1999).

In summary, we have shown that the contribution of the MHC to genetic risk for RA involves two genetic regions in addition to the DRB1 locus. These results emphasize the utility of using haplotypic analysis, instead of single-marker analysis, to dissect these genetic effects. At the same time, the extensive linkage disequilibrium within the human MHC (Jeffreys et al. 2001) will make it difficult to identify precisely which gene or genes are responsible for these associations. It is possible that larger sample sizes may lead to the identification of ancestral recombinants that will allow further narrowing of these critical regions. This will also provide adequate statistical power to define interactions between these risk factors, which appear to exist in DRB1*0401/*0404 compound heterozygotes and can now be addressed with respect to the 8.1 ancestral haplotype. Nevertheless, the identification of the risk genes in these regions will be challenging and, ultimately, will depend on relating candidate polymorphisms to the specific biological alterations that are involved in disease pathogenesis.

Acknowledgments

This work was supported by National Institutes of Health grants R01-AR44222 and N01-AR7-2232 (to P.K.G.) and R01-AR43274 and P01-AR45231-SCOR in SLE (to T.W.B.). These studies were performed, in part, in the General Clinical Research Center, Moffitt Hospital, University of California, San Francisco, with funds provided by National Center for Research Resources grant 5-M01-RR-00079. The NARAC includes a subset of the authors (P.K.G., L.A.C., M.F.S., and C.I.A.), as well as the following individuals: Salvatore Albani (University of California, San Diego), S. Louis Bridges (University of Alabama at Birmingham), Leigh F. Callahan (University of North Carolina), Daniel O. Clegg (University of Utah), Donald Flemming (Bethesda Naval Medical Center), Daniel L. Kastner (National Institute of Arthritis and Musculoskeletal and Skin Diseases), Marlene Kern (North Shore–Long Island Jewish Research Institute), J. Lee Nelson (Fred Hutchinson Cancer Research Center), Theodore Pincus (Vanderbilt University), David S. Pisetsky (Duke University), Richard Pope (Northwestern University), Harry W. Schroeder (University of Alabama at Birmingham), Ryk Ward (Oxford University), and Mark H. Wener (University of Washington). Support for the NARAC is provided, in part, by the National Arthritis Foundation. Finally, we wish to thank the patients and families who made these studies possible through their generous participation.

Electronic-Database Information

Accession numbers and URLs for data presented herein are as follows:

Center for Medical Genetics, Marshfield Medical Research Foundation, http://research.marshfieldclinic.org/genetics/Lab_Methods/methods.htm

Immunogenetics/HLA Sequence Satabase, http://www.ebi.ac.uk:80/imgt/hla/nomen.html

North American Rheumatoid Arthritis Consortium, http://www.naracdata.org/

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for RA [180300], common variable immunodeficiency [240500], IgA deficiency [137100], systemic lupus erythematosus [152700], insulin-dependent diabetes mellitus [222100], autoimmune hepatitis [142395], and Epstein-Barr–virus disease [226990]

Sanger Institute, http://www.sanger.ac.uk/HGP/Chr6/MHC.shtml (for HLA consensus sequence)

UCSC Genome Bioinformatics, http://www.genome.ucsc.edu (for exact marker locations)

References

Alper CA, Kruskall MS, Marcus-Bagley D, Craven DE, Katz AJ, Brink SJ, Dienstag JL, Awdeh Z, Yunis EJ (1989) Genetic

prediction of nonresponse to hepatitis B vaccine. N Engl J Med 321:708–712

Candore G, Cigna D, Gervasi F, Colucci AT, Modica MA, Caruso C (1994) In vitro cytokine production by HLA-B8,DR3 positive subjects. Autoimmunity 18:121–132

Caruso C, Candore G, Colucci AT, Cigna D, Sammartano F, Ammatuna P (1993) HLA-B8,DR3 phenotype and the antibody response against Epstein-Barr virus. Immunol Invest 22:41–51

Cornelis F, Faure S, Martinez M, Prud'homme JF, Fritz P, Dib C, Alves H, et al (1998) New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. Proc Natl Acad Sci USA 95:10746–10750

Deighton CM, Walker DJ, Griffiths ID, Roberts DF (1989) The contribution of HLA to rheumatoid arthritis. Clin Genet 36:178–182

Dizier MH, Eliaou JF, Babron MC, Combe B, Sany J, Clot J, Clerget-Darpoux F (1993) Investigation of the HLA component involved in rheumatoid arthritis (RA) by using the marker association-segregation χ^2 (MASC) method: rejection of the unifying-shared-epitope hypothesis. Am J Hum Genet 53:715–721

Erlich HA, Bugawan T, Begovich AB, Scharf S, Griffith R, Saiki R, Higuchi R, Walsh PS (1991) HLA-DR, DQ and DP typing using PCR amplification and immobilized probes. Eur J Immunogenet 18:33–55

Graham RR, Ortmann WA, Langefeld CD, Jawaheer D, Selby SA, Rodine PR, Baechler EC, Rohlf KE, Shark KB, Nair RP, Stuart PE, Elder JT, King RA, Moser KL, Gaffney PM, Bugawan TL, Erlich HA, Rich SS, Gregersen PK, Behrens TW (2002) Visualizing human leukocyte antigen class II risk haplotypes in human systemic lupus erythematosus. Am J Hum Genet 71:543–553 (in this issue)

Gregersen PK (1998) The North American Rheumatoid Arthritis Consortium: bringing genetic analysis to bear on disease susceptibility, severity, and outcome. Arthritis Care Res 11:1–2

Gregersen PK, Silver J, Winchester RJ (1987) The shared epitope hypothesis: an approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum 30:1205–1213

Hall FC, Weeks DE, Camilleri JP, Williams LA, Amos N, Darke C, Gibson K, Pile K, Wordsworth BP, Jessop JD (1996) Influence of the HLA-DRB1 locus on susceptibility and severity in rheumatoid arthritis. QJM 89:821–829

James JA, Neas BR, Moser KL, Hall T, Bruner GR, Sestak AL, Harley JB (2001) Systemic lupus erythematosus in adults is associated with previous Epstein-Barr virus exposure. Arthritis Rheum 44:1122–1126

Jawaheer D, Seldin MF, Amos CI, Chen WV, Shigeta R, Monteiro J, Kern M, Criswell LA, Albani S, Nelson JL, Clegg DO, Pope R, Schroeder HW Jr, Bridges SL Jr, Pisetsky DS, Ward R, Kastner DL, Wilder RL, Pincus T, Callahan LF, Flemming D, Wener MH, Gregersen PK (2001) A genome-wide screen in multiplex rheumatoid arthritis families suggests genetic overlap with other autoimmune diseases. Am J Hum Genet 68:927–936

Jeffreys AJ, Kauppi L, Neumann R (2001) Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. Nat Genet 29:217–222

Kaslow RA, Duquesnoy R, VanRaden M, Kingsley L, Marrari M, Friedman H, Su S, et al (1990) A1, Cw7, B8, DR3 HLA antigen combination associated with rapid decline of T-helper lymphocytes in HIV-1 infection: a report from the Multicenter AIDS Cohort Study. Lancet 335:927–930

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet 58:1347–1363

Lio D, Candore G, Colombo A, Colonna Romano G, Gervasi F, Marino V, Scola L, Caruso C (2001) A genetically determined high setting of TNF-alpha influences immunologic parameters of HLA-B8,DR3 positive subjects: implications for autoimmunity. Hum Immunol 62:705–713

Martin ER, Monks SA, Warren LL, Kaplan NL (2000) A test for linkage and association in general pedigrees: the pedigree disequilibrium test. Am J Hum Genet 67:146–154

McDaniel DO, Alarcon GS, Pratt PW, Reveille JD (1995) Most African-American patients with rheumatoid arthritis do not have the rheumatoid antigenic determinant (epitope). Ann Intern Med 123:181–187

MCH Sequencing Consortium (1999) Complete sequence and gene map of a human major histocompatibility complex. Nature 401:921–923

Milner CM, Campbell RD (2001) Genetic organization of the human MHC class III region. Front Biosci Suppl 6: D914–D926

Moxley G, Cohen HJ (2002) Genetic studies, clinical heterogeneity, and disease outcome studies in rheumatoid arthritis. Rheum Dis Clin North Am 28:39–58

Mulcahy B, Waldron-Lynch F, McDermott MF, Adams C, Amos CI, Zhu DK, Ward RH, Clegg DO, Shanahan F, Molloy MG, O'Gara F (1996) Genetic variability in the tumor necrosis factor-lymphotoxin region influences susceptibility to rheumatoid arthritis. Am J Hum Genet 59:676–683

Nair RP, Stuart P, Henseler T, Jenisch S, Chia NV, Westphal E, Schork NJ, Kim J, Lim HW, Christophers E, Voorhees JJ, Elder JT (2000) Localization of psoriasis-susceptibility locus PSORS1 to a 60-kb interval telomeric to HLA-C. Am J Hum Genet 66:1833–1844

Nepom GT, Seyfried CE, Holbeck SL, Wilske KR, Nepom BS (1986) Identification of HLA-Dw14 genes in DR4+ rheumatoid arthritis. Lancet 2:1002–1005

Ollier W, Thomson W (1992) Population genetics of rheumatoid arthritis. Rheum Dis Clin North Am 18:741–759

Ota M, Katsuyama Y, Kimura A, Tsuchiya K, Kondo M, Naruse T, Mizuki N, Itoh K, Sasazuki T, Inoko H (2001) A second susceptibility gene for developing rheumatoid arthritis in the human MHC is localized within a 70-kb interval telomeric of the TNF genes in the HLA class III region. Genomics 71:263–270.

Price P, Witt C, Allcock R, Sayer D, Garlepp M, Kok CC, French M, Mallal S, Christiansen F (1999) The genetic basis for the association of the 8.1 ancestral haplotype (A1, B8, DR3) with multiple immunopathological diseases. Immunol Rev 167:257–274

Ronningen KS, Ploski R, Hansen T, Thorsby E (1992) Dw14 is a Dw4-independent risk factor for rheumatoid arthritis among Norwegians. Tissue Antigens 39:280

Seldin MF, Amos CI, Ward R, Gregersen PK (1999) The ge-

netics revolution and the assault on rheumatoid arthritis. Arthritis Rheum 42:1071–1079

Shannon CE (1948) The mathematical theory of communication. Bell Syst Tech J 27:623–656

Sobel E, Lange K (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet 58:1323–1337

Stastny P (1978) Association of the B-cell alloantigen DRw4 with rheumatoid arthritis. N Engl J Med 298:869–871

Teller K, Budhai L, Zhang M, Haramati N, Keiser HD, Davidson A (1996) HLA-DRB1 and DQB typing of Hispanic American patients with rheumatoid arthritis: the "shared epitope" hypothesis may not apply. J Rheumatol 23:1363–1368

Thomson W, Harrison B, Ollier B, Wiles N, Payton T, Barrett J, Symmons D, Silman A (1999) Quantifying the exact role of HLA-DRB1 alleles in susceptibility to inflammatory polyarthritis: results from a large, population-based study. Arthritis Rheum 42:757–762

Tosato G, Steinberg AD, Blaese RM (1981) Defective EBV-specific suppressor T-cell function in rheumatoid arthritis. N Engl J Med 305:1238–1243

Udalova IA, Nedospasov SA, Webb GC, Chaplin DD, Turetskaya RL (1993) Highly informative typing of the human TNF locus using six adjacent polymorphic markers. Genomics 16:180–186

Waldron-Lynch F, Adams C, Amos C, Zhu DK, McDermott MF, Shanahan F, Molloy MG, O'Gara F (2001) Tumour necrosis factor 5′ promoter single nucleotide polymorphisms influence susceptibility to rheumatoid arthritis (RA) in immunogenetically defined multiplex RA families. Genes Immun 2: 82–87

Wilson AG, Symons JA, McDowell TL, McDevitt HO, Duff GW (1997) Effects of a polymorphism in the human tumor necrosis factor alpha promoter on transcriptional activation. Proc Natl Acad Sci USA 94:3195–3199

Yen JH, Moore BE, Nakajima T, Scholl D, Schaid DJ, Weyand CM, Goronzy JJ (2001) Major histocompatibility complex class I-recognizing receptors are disease risk genes in rheumatoid arthritis. J Exp Med 193:1159–1167

Zanelli E, Jones G, Pascual M, Eerligh P, van der Slik AR, Zwinderman AH, Verduyn W, Schreuder GM, Roovers E, Breedveld FC, de Vries RR, Martin J, Giphart MJ (2001) The telomeric part of the HLA region predisposes to rheumatoid arthritis independently of the class II loci. Hum Immunol 62:75–84