Fine-Scale Mapping at IGAD1 and Genome-Wide Genetic Linkage Analysis Implicate HLA-DQ/DR as a Major Susceptibility Locus in Selective IgA Deficiency and Common Variable Immunodeficiency

Jana Kralovicova, Lennart Hammarström, Alessandro Plebani, A. David B. Webster and Igor Vorechovsky

*J Immunol* 2003; 170:2765-2775; doi: 10.4049/jimmunol.170.5.2765

http://www.jimmunol.org/content/170/5/2765

**Supplementary Material**

http://www.jimmunol.org/content/suppl/2003/03/07/170.5.2765.DC2

**References**

This article cites 86 articles, 18 of which you can access for free at:

http://www.jimmunol.org/content/170/5/2765.full#ref-list-1

**Subscription**

Information about subscribing to *The Journal of Immunology* is online at:

http://jimmunol.org/subscription

**Permissions**

Submit copyright permission requests at:

http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**

Receive free email-alerts when new articles cite this article. Sign up at:

http://jimmunol.org/alerts
Fine-Scale Mapping at IGAD1 and Genome-Wide Genetic Linkage Analysis Implicate HLA-DQ/DR as a Major Susceptibility Locus in Selective IgA Deficiency and Common Variable Immunodeficiency

Jana Kralovicova,* Lennart Hammarström, † Alessandro Plebani,‡ A. David B. Webster,§ and Igor Vorechovsky•2*

Selective IgA deficiency (IgAD) and common variable immunodeficiency (CVID) are the most common primary immunodeficiencies in humans. A high degree of familial clustering, marked differences in the population prevalence among ethnic groups, association of IgAD and CVID in families, and a predominant inheritance pattern in multiple-case pedigrees have suggested a strong, shared genetic predisposition. Previous genetic linkage, case-control, and family-based association studies mapped an IgAD/CVID susceptibility locus, designated IGAD1, to the MHC, but its precise location within the MHC has been controversial. We have analyzed a sample of 101 multiple- and 110 single-case families using 36 markers at the IGAD1 candidate region and mapped homozygous stretches across the MHC shared by affected family members. Haplotype analysis, linkage disequilibrium, and homozygosity mapping indicated that HLA-DQ/DR is the major IGAD1 locus, strongly suggesting the autoimmune pathogenesis of IgAD/CVID. This is supported by the highest excess of allelic sharing at 6p in the genome-wide linkage analysis of 101 IgAD/CVID families using 383 marker loci, by previously reported restrictions of the T cell repertoires in CVID, the presence of autoantibodies, impaired T cell activation, and a dysregulation of a number of genes in the targeted immune system. IgAD/CVID may thus provide a useful model for the study of pathogenesis and novel therapeutic strategies in autoimmune diseases. The Journal of Immunology, 2003, 170: 2765–2775.

AAlthough a number of disease-causing mutations have been identified in monogenic primary immunodeficiencies (PIDs) in the last 15 years (1), the molecular mechanisms underlying the development of selective IgA deficiency (IgAD) and common variable immunodeficiency (CVID), the most common PIDs in Caucasoid populations, have not been elucidated. Both PIDs have been characterized by a defect in the terminal stage of lymphocyte differentiation, leading to an impaired production of one or more Ig isotypes, resulting in susceptibility to infections (1, 2). Clinically overt CVID necessitates long-term Ig replacement and antimicrobial therapy with considerable health-care expenditures, but treated patients still show a higher mortality than the general population (3).

Although IgAD/CVID do not show Mendelian segregation in families, they exhibit a high degree of familial clustering (4, 5) and marked differences in the population prevalence among ethnic groups, ranging from ~1/500 in Caucasians to 1/18500 in Japanese (for review see Ref. 2), supporting a strong involvement of hereditary factors. The association of both PIDs in families (6, 7), individuals (8–12), and with identical HLA Ags (6, 13, 14) strongly suggested that the two clinically discernible disorders share genetic predisposition. Screening of family members of index cases for serum Ig levels showed that about a third of multiply affected (multiplex) families had both CVID and IgAD in close relatives, typically CVID in the parental generation followed by IgAD in the offspring (7). These observations indicated that a large subset of cases diagnosed as CVID, which is much less prevalent than IgAD (~1/25,000 in the U.K.; A. D. B. Webster, unpublished data), represents a more severe manifestation of a defect in common.

Functional lymphocyte studies in CVID have revealed a number of abnormalities (for review see Ref. 15), but the elucidation of a primary defect has been hampered by the complexities of interactions between immune effector cells. This led us to initiate a genetic linkage study to identify chromosome susceptibility loci. We previously found a significant increase of allele sharing on chromosome 6p (7), consistent with MHC associations in case-control (for review see 2) and family-based (7) association studies, providing the proof for the presence of a susceptibility locus designated IGAD1. This evidence was further supported by a parental allele segregation distortion at IGAD1 accompanied by parent-of-origin penetrance differences of IgAD, and by the observation of a higher frequency of anti-IgA Abs (Ab) in females transmitting the disease to the offspring than in female nontransmitters, implicating...
anti-IgA Ab in familial clustering (7). The evidence for genetic linkage to *IGAD1* was stronger in families with anti-IgA-positive female transmitters than in pedigrees with anti-IgA-negative individuals, suggesting a major role of *IGAD1* in parent-of-origin penetrance effects and autoantibody production (7). However, the relative significance of this locus in the overall predisposition to IgAD/CVID has not been clear. Likewise, there has been no consensus as to the precise location of *IGAD1*, with the disease locus placed to the *HLA* class I/III regions (6, 16–19) or more centromerically to the *HLA* class II region (20, 21). Could this decade-lasting controversy (16, 20, 22) be resolved using genetic means? In the present study, we have typed an extensive set of families with IgAD/CVID using MHC short tandem repeats (STR) and single nucleotide polymorphisms (SNPs), constructed haplotypes in the region, and monitored family-based allelic associations in the MHC. In addition, we have determined *HLA* specificities in representative carriers of STR haplotypes and provide in this report a reference panel of STR haplotypes/alleles on the most prevalent *HLA* class II haplotypes in a Caucasian population. We also show a significant risk of developing IgAD/CVID conferred by homozygous stretches in the class II region. A maximum excess of the observed homozygosity over expected under the Hardy-Weinberg equilibrium (HWE) was in the region identical to that implicated by LD fine-mapping methods, with *HLA-DQ/DR* as the only expressed genes. In addition, the genome-wide linkage analysis of 101 multiplex families with IgAD/CVID showed the highest excess of allele sharing at 6p and suggested putative non-MHC loci at 4p, 12p, and 14q. Finally, we propose the autoimmune pathogenesis of IgAD/CVID and discuss how the previously observed lymphocyte abnormalities in CVID can be explained in the context of peptide-DQ/DR interactions with TCR.

**Materials and Methods**

**Ascertainment of probands and families**

Diagnosis of patients with IgAD and CVID was established in accordance with accepted recommendations as reported earlier (7, 23). The affection status and the ascertainment of index cases were defined previously (7). Multiplex families were identified through proband by screening blood relatives of index cases for serum Ig levels by nephelometry (7). The measurements were conducted blind to family relationship and affection status. The ethnic origin of multiplex families was shown previously (21). The Ig measurements of samples from outside Sweden were repeated in the Swedish laboratory to guarantee that the same method defined the phenotype and to ensure that a single blood sample could serve both for the determination of the affection status and DNA extraction, decreasing the probability of a laboratory mix-up. Bilineal families, rare cases with previously identified deletions in the Ig genes (24), drug-induced IgAD, and two multiplex families with IgAD in which the phenotype could not be regarded as categorical, were excluded from the study. Although no multiple-case family showed male-to-male inheritance of CVID, suggesting an X-linked defect, CVID samples were analyzed for mutations in *RTK* (25), *AIF* (J. Kralovicova and I. Vorechovsky, unpublished observations), and *HIGM1* (A. D. B. Webster et al., unpublished observations), and *HIGM1* (25), with the disease locus shown previously for the analyzed sample (28). For the genome-wide scan, we used a total of 383 STR loci with an average intermarker genetic distance of 11 cM. Oligonucleotide primers and PCR conditions, Mg2+ concentrations, and annealing temperatures are shown as supplementary information. In addition, genotypes at three CSNPs in the **BTLN2** gene (939A→G, 1050G→A, and 1078A→G), located ~20 kb centromeric of **LHI** (Fig. 1) (29), were determined using nucleotide sequencing with BigDye terminators (Applera, Norwalk, CT) as described (26) to establish a telomeric limit of the candidate region on a subset of haplotypes. The amplification primers were 5′-CCC CAC CTC ACC TAA G-3′ and 5′-AGA GAA ATT GTC CAG GAA CTA-3′.

**Linkage analysis and linkage disequilibrium (LD) mapping**

The Genehunter (version 2.0) (30) and Allegro (version 1.1) (31) programs were used for computing NPL and Zlr scores, respectively. Haplotypes were constructed using the Genehunter and Simwalk2 algorithms. *HLA* haplotypes were inspected and corrected manually against the Genotyper plots. Samples from pedigrees with variant or inconsistent alleles on the major susceptibility haplotypes were retyped to confirm the data.

**Risk calculations and HWE testing**

Population-attributable risk, or a fraction of affected cases in the studied population that would have not occurred had the risk factor been absent, was estimated as \( \text{f(RR-1)/RR} \) (34), where \( f \) is the proportion of cases carrying the risk haplotype and RR is the relative risk of developing IgAD/CVID.

The excess of homozygosity over expected was calculated from the observed \( (H_o) \) and expected \( (H_e) \) values in affected individuals as a parameter \( F = (H_o - H_e)(1 - H_e) \) (35). \( H_o \) was computed from allelic frequencies assuming HWE.

**Mutation analysis and sequence-based HLA typing**

PCR single-strand conformation polymorphism analysis was used for mutation detection in non-MHC genes as reported previously (26). Oligonucleotide primer sequences for PCR amplification are shown as supplementary information. Two single-strand conformation polymorphism gel-running conditions were used for the ICOS and BST1 screening. HLA specificities were determined using sequence-based typing as described (28) in representative carriers of STR haplotypes to relate HLA specificities and STR haplotypes in close proximity of the *HLA* class II genes.

**Results**

**Family-based associations and fine mapping at *IGAD1***

To narrow down the *IGAD1* candidate region, we genotyped 879 individuals from 210 families at 36 *HLA* STR loci selected from a large number of *HLA* STRs for correct Mendelian inheritance, reliable size clustering, and the absence of null alleles. We first determined the degree of family-based allelic association with IgAD/
CVID using the FBAT algorithm (Fig. 1). We found the most significant associations in the telomeric part of the class II region, decreasing rapidly in the area of known recombination hot spots centromeric of \( HLA-DQB1 \), but less precipitously in the opposite direction (Fig. 1), partly reflecting the progression of LD in the region (28, 36). These results illustrated that a mere degree of association does not permit a safe placement of the disease gene to a smaller area in this high LD, selection-driven, and gene-rich region, prompting us to study haplotypes in pedigrees.

We next used the trimmed haplotype method using the analysis of 2-, 3-, and 4-locus STR haplotypes that could have descended from the same ancestral founder but have been trimmed in succeeding generations (33). We show these results as a comparison of two \( IGAD1 \) subregions (Table I). Region 1 contains genes encoding \( HLA-DQ \) and \(-DR\) molecules, whereas region 2, suggested in previous studies (6, 16–19), is in the telomeric part of the class III region and spans a similar physical distance (Fig. 1). We found that the region 1 haplotypes were always ranked at the top of trimmed haplotype tables with highly significant likelihood ratio/regression/model-dependent statistics and constituted the majority of the 5% best performing ancestral haplotypes tested (Table I). The interval between \( G51152 \) and \( 9-99431 \), which contains only \( HLA-DQB1, -DQA1, -DRB1, \) and \( -DRA (-DRB3 \) is present on \( HLA-DQB1*0201, DRB1*0301 \) haplotypes) and few \( HLA \) pseudogenes (Fig. 1), was implicated by all Trimhap routines as the best candidate, including model-dependent algorithms with varying program parameters (Table I). For 2-locus haplotypes, the regression analysis indicated the \( DQCAR-DQCARII \) interval telomeric of \( DQB1 \) as the most likely location (Fig. 1), whereas statistics 1 and 3 favored the region between \( G51152-DQCAR \), containing only \( DQB1 \), suggesting limitations of

**FIGURE 1.** Family-based associations at the \( IGAD1 \) candidate region in 242 nuclear families. The map is drawn to scale, reflecting a particular haplotype sequenced (for sequence, see supplementary information). Only a subset of MHC genes relative to tested STRs is shown. The STRs used for allelic association are in bold as they were found to show correct Mendelian inheritance, unambiguous clustering and scoring, and the absence of null alleles (28). The FBAT program (32) was used under the additive and dominant models in biallelic and multiallelic tests. The empirical variance option of FBAT was used for computing Z scores. In all tests, additive models performed better for most marker loci than dominant models.
An ancestral mutation event, mutation rate per generation of 10
by bootstrap methods (33) are shown for the top 10 haplotypes. The following program parameters were used in the model-dependent analysis (statistics 3): 200 generations since
230 kb) in the border area between the
class II region and region 2 is between 24-140297 and MIB (~240 kb) in the border area between the HLA class III and class I regions (Fig. 1). The numbers indicate the proportion (percentage in brackets) of n-locus haplotypes encompassing at least one STR in the tested region among 5% of haplotypes ranked at the top of the trimmed haplotype table. The corresponding statistics and p-values generated by bootstrap methods (33) are shown for the top 10 haplotypes. The following program parameters were used in the model-dependent analysis (statistics 3): 200 generations since an ancestral mutation event, mutation rate per generation of 10^{-8}, genotyping error rate of 0.01, one test-point per marker-marker subinterval, and 1000 replicate samples generated by permutation bootstrapping. The proportion of haplotypes descended from a given ancestral founder haplotype was set to 0.2.

The method for such small regions. The HLA class I haplotypes were at the bottom of trimmed haplotype tables, consistent with weaker associations in TDTs (data not shown), thus, not supporting the presence of a major disease gene. Consistent with the trimmed haplotype analysis, the extended TDT (37) using the same data set indicated that the most likely location of the disease gene is in the HLA-DQ/DR region between STRs G51152 and LHI (Figs. 1 and 2).

Homozygosity at HLA-DQ/DR

Because previous association studies suggested an increase of homozygosity in affected individuals vs controls both for HLA-DQ/DR (38, 39) and HLA STRs (21), we estimated the relative risk of IgAD/CVID conferred by homozygous stretches at adjacent STRs and examined which region was associated with the highest risk. We chose to test 4- and 5-STR sliding windows because the average probability of being homozygous by chance was sufficiently low (~2 × 10^{-5} and 3 × 10^{-5}, respectively), minimizing random homozygosity at adjacent STRs. Among 210 probands, we found 34 (16.2%) cases (17 sporadic and 17 familial; 15 CVID and 19 IgAD) homozygous at four or more adjacent marker loci compared with only 7.7% (17/220) in unaffected founder controls (λ^2 = 7.3, p < 0.01, odds ratio = 2.3, 95% confidence interval 1.3–4.2), indicating that homozygosity at adjacent STRs at IGAD1 confers susceptibility. We also identified 46 such defined homozygotes among a total of 258 affected individuals in multiplex families (17.8%) as compared with only 22 in 231 (9.5%) nonaffected relatives, consistent with a higher number of susceptibility haplotypes segregating in related family members.

We next examined the distribution of homozygous regions across IGAD1. We found that the interval between G51152 and 9-99431 (region 1), for which the expected homozygosity was only 6 × 10^{-4}, was encompassed in 31/34 index cases homozygous in at least four adjacent HLA STRs. In contrast, only three probands showed such homozygosity in the HLA class III region and not in the HLA-DQ/DR region, with only two of them between 24-140297 and MIB (region 2, Fig. 1). Ten of 34 probands were homozygous for the whole tested region. We also analyzed all affected homozygotes in families and found a similar bias toward the G51152/9-99431 interval (Fig. 3a, b, d, and f), thus strongly implicating region 1 as containing disease gene(s).

For the five-marker sliding window, the frequencies of homozygous index cases vs controls were 13.1 and 5.9% (p < 0.01). Of 26 homozygous probands, only a single case exhibited homozygosity in the telomeric part of the class III region and not in the HLA-DQ/DR region. The frequencies of unaffected homozygous individuals (7.7% for 4-STRs and 5.9% for 5-STRs) were comparable to ~6–8% homozygosity observed in large numbers of population controls typed for HLA-DR or -DQ specificities (38, 39).

To exclude that the excess of homozygosity in the class II region was due to other factors than the segregation of disease alleles in affected family members, we constructed HLA STR haplotypes for 128 control “child-father-mother” trios ascertained through Swedish children with cystic fibrosis and phenylketonuria. In contrast to affected individuals, we found no bias in the distribution of homozygous stretches toward the HLA class II region. All haplotypes in IgAD/CVID families and control trios are shown as supplementary information online.4

If there is evidence for a significant increase of homozygosity in affected individuals as compared with unaffected controls at IGAD1 and if there is a support for an additive disease model (Fig. 1), then testing for deviation from the HWE may be helpful in fine

Table 1. Trimmed haplotype analysis of 210 families with IgAD/CVID

<table>
<thead>
<tr>
<th>No. of Tested Founder Haplotypes</th>
<th>Number of Loci in Ancestral Haplotype</th>
<th>Region 1</th>
<th>Region 2</th>
<th>Range for Top 10</th>
<th>Region 1</th>
<th>Region 2</th>
<th>Range for Top 10</th>
<th>Region 1</th>
<th>Region 2</th>
<th>Range for Top 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>266</td>
<td>2</td>
<td>10/13(77)</td>
<td>2/13(15)</td>
<td>13.0–7.9 (p &lt; 10^{-7})</td>
<td>9/13(69)</td>
<td>1/13(8)</td>
<td>3.03–1.35–4</td>
<td>10/13(77)</td>
<td>1/13(8)</td>
<td>5.9–3.8 (p &lt; 10^{-4})</td>
</tr>
<tr>
<td>1926</td>
<td>3</td>
<td>85/96(89)</td>
<td>5/96(5)</td>
<td>14.2–11.5 (p &lt; 10^{-7})</td>
<td>74/96(77)</td>
<td>21/96(22)</td>
<td>6.75–4.8–7</td>
<td>72/96(75)</td>
<td>13/96(14)</td>
<td>6.7–5.4 (p &lt; 10^{-4})</td>
</tr>
<tr>
<td>10218</td>
<td>4</td>
<td>483/511(95)</td>
<td>23/511(5)</td>
<td>14.5–13.4 (p &lt; 10^{-7})</td>
<td>432/511(85)</td>
<td>76/511(15)</td>
<td>1.20–6.17–4</td>
<td>405/511(79)</td>
<td>93/511(18)</td>
<td>8.4–6.3 (p &lt; 10^{-4})</td>
</tr>
</tbody>
</table>

All pedigree founder haplotypes were sorted in a trimmed haplotype table with the ancestral haplotype at its head and mutually exclusive set of other haplotypes in decreasing order of similarity to the ancestral haplotype (33). Region 1 is between G51152 and 9-99431 (~240 kb) in the HLA class II region and region 2 is between 24-140297 and MIB (~230 kb) in the border area between the HLA class III and class I regions (Fig. 1). The numbers indicate the proportion (percentage in brackets) of n-locus haplotypes encompassing at least one STR in the tested region among 5% of haplotypes ranked at the top of the trimmed haplotype table. The corresponding statistics and p-values generated by bootstrap methods (33) are shown for the top 10 haplotypes. The following program parameters were used in the model-dependent analysis (statistics 3): 200 generations since an ancestral mutation event, mutation rate per generation of 10^{-8}, genotyping error rate of 0.01, one test-point per marker-marker subinterval, and 1000 replicate samples generated by permutation bootstrapping. The proportion of haplotypes descended from a given ancestral founder haplotype was set to 0.2.

FIGURE 2. Fine-scale mapping of IGAD1. The extended TDT (A) and a proportional excess of homozygosity (F) (B) in the HLA class II and class III regions.
localization of disease genes in traits with presumed genetic heterogeneity (35, 40). To measure this, we calculated the excess of homozygosity over expected in affected individuals (Fig. 2B). We observed the largest proportionate increase of homozygosity (F) at 9–99431, with the Fmax value between this marker and DQCAII, implicating the same location as by LD and haplotype analyses, and providing independent support for HLA-DQ/DR as the disease susceptibility locus.

We then determined the nucleotide sequence of gene-coding and promoter regions of HLA-DRB3 (present only on ancestral haplotypes AHR.I), -DQB1, -DQA1, -DRB1, and -DRA in affected homozygotes. As expected, we observed no sequence changes as compared with the published HLA polymorphisms, consistent with our proposal of the autoimmune pathogenesis of IgAD/CVID.

**STR vs HLA haplotypes**

To relate STR haplotypes to HLA specificities, we determined the nucleotide sequence of exons 2 and 3 of HLA-DQB1 and -DRB1 in 86 individuals carrying representative STR haplotypes between G51152 and 9–99431 (Table II). We identified only 11 (6.4%) variant alleles, indicating that the majority of tested STR haplotypes carried identical high-resolution HLA-DQB1, -DRB1 haplotypes. Only a single allelic variant was found at HLA-DQB1 for the most predisposing haplotypes, suggesting a very high predictive value of STR haplotyping for deducing HLA-DQ/DR specificities. Using multipoint TDTs, we ordered STR haplotypes from the most disease predisposing to the most protective (Table II). The most protective haplotype HLA-DQB1*0602, DRB1*15011 was observed in trans with the majority of STR haplotypes in affected individuals, except for those carrying DQB1*06 (Table II). The absence of deduced DQB1*0602/DQB1*06xx genotypes among 367 affected individuals suggests that DQB1*0602 homozygosity may confer absolute protection from IgAD/CVID and indicates the existence of a dosage effect both for protective and susceptibility HLA-DQ/DR haplotypes. In addition, we found no obvious preference for the susceptibility/protective haplotypes to occur in cis/trans combination with each other in affected individuals, consistent with a less stringent HLA restriction than that found, for example, in celiac disease, which occurs almost exclusively on the background of the HLA-DQB1*0201 alleles.

Because the HLA-DQB1*0201, -DRB1*03011 haplotypes could be unambiguously predicted using STR haplotypes between

![Figure 3](http://www.jimmunol.org/)

**FIGURE 3.** Alignment of susceptibility haplotypes in affected family members. Shared haplotype portions are highlighted in gray (haplotypes extending into the class I region) or blue (haplotypes terminating in the class III/I regions). Variant alleles/haplotypes are in light background colors. Haplotypes were constructed using the Genehunter program (version 2.0) and edited manually after retyping. Family designation starting with cv denotes a multiplex family, whereas the remaining cases were sporadic. CVID patients are boxed; the remaining family members had IgAD. We show only homozygous carriers of the most prevalent susceptibility haplotypes (a and d) and both homozygous and heterozygous carriers for less-prevalent haplotypes (b, c, e, and f). Full haplotypes in each family member are available as supplementary information online.
Table II. Sequence-based HLA typing of representative STR haplotypes

<table>
<thead>
<tr>
<th>Number of Analyzed STR Haplotypes</th>
<th>Number of STR Haplotypes with Identical HLA Alleles</th>
<th>Predominant HLA Haplotypes</th>
<th>Variant HLA Allele(s)</th>
<th>STR Haplotypes&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Multipoint Transmission Disequilibrium Test for 2-, 3- and 4-locus STR Haplotypes (T/NT, &lt;sup&gt;b&lt;/sup&gt; p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>28</td>
<td>0201 0301</td>
<td>—</td>
<td>5 1 7 6</td>
<td>35/24 37/19 82/38 30/18 35/18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>05011 0101</td>
<td>DQBI*0103</td>
<td>10 3 6 8</td>
<td>0.15 0.02 6.10&lt;sup&gt;-5&lt;/sup&gt; 0.08 0.02 0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>05011 0102</td>
<td>—</td>
<td>9 3 6 2</td>
<td>0.005 10&lt;sup&gt;-4&lt;/sup&gt; 0.001 0.005 0.004 0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>05011 1001</td>
<td>—</td>
<td>9 3 4 5</td>
<td>0.003 10&lt;sup&gt;-4&lt;/sup&gt; 0.06 0.008 0.02 0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>05031 1401 1404</td>
<td>DRBI*1404</td>
<td>6 5 5 5A2</td>
<td>18/4 5/2 12/13 5/2 5/2 5/2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0202 0701</td>
<td>—</td>
<td>4 11 14 2</td>
<td>0.003 0.004 NS 0.007 0.01 0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>03011 11011 1104</td>
<td>DRBI*1104</td>
<td>4 11 2 5</td>
<td>22/11 9/0 22/21 9/0 9/0 9/0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0302 0401</td>
<td>DRBI*0401&lt;sup&gt;c&lt;/sup&gt;</td>
<td>9/4 6 3 6/7</td>
<td>0.05 0.003 NS 0.003 0.003 0.003</td>
</tr>
<tr>
<td>15</td>
<td>11</td>
<td>0302 0401</td>
<td>DRBI*0401&lt;sup&gt;c&lt;/sup&gt;</td>
<td>9/4 6 3 6/7</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0302 0403</td>
<td>—</td>
<td>9 6 5 6</td>
<td>NS NS NS NS NS NS NS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>06041 1302</td>
<td>—</td>
<td>7 3 10 5</td>
<td>NS NS NS NS NS NS NS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0302 0701</td>
<td>—</td>
<td>4 10 12/11 1</td>
<td>NS NS NS NS NS NS NS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0302 0901</td>
<td>DQBI*0902</td>
<td>4 8 4 5</td>
<td>NS NS NS NS NS NS NS</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>0202 0701</td>
<td>—</td>
<td>4 11 14 2</td>
<td>0.003 0.004 NS 0.007 0.01 0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>03011 1201</td>
<td>—</td>
<td>4 9 1 2</td>
<td>0.007 0.02 NS 0.02 0.06 0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0603 1301</td>
<td>—</td>
<td>8 3 9 5</td>
<td>5/18 2/10 7/11 2/10 2.8 2.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0603 1301 1302</td>
<td>DRBI*1302</td>
<td>8 3 9 4</td>
<td>12/40 10/15 7/4 8/11 3/2 3/2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0602 15011</td>
<td>DQBI*0611</td>
<td>8 3 9 5</td>
<td>12/40 10/15 12/16 8/11 6/11 4/7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0602 15011</td>
<td>DQBI*0611</td>
<td>8 3 5 1</td>
<td>0.001 0.003 0.001 0.003 0.001 0.002</td>
</tr>
</tbody>
</table>

<sup>a</sup> HLA-DQB1 is located between G51152 and DQCAR and HLA-DRB1 is between DQCARII and 9-99441 (Fig. 1).

<sup>b</sup> T(NT), No. of (non)transmissions from heterozygous parents. The most predisposing haplotypes are at the top and the most protective haplotypes are at the bottom of the table.

<sup>c</sup> Variant alleles were DRBI*0403, *0404 and *0410. The Genehunter program (version 2) was used for multipoint TDT.
part of the class II region and had variant haplotypes in the class C. To map the telomeric border of this predisposing haplotype, 3 telomeric class III alleles and only 5 of them class I alleles (Fig. 3). The affected 10-3-6-8 homozygotes, only 9 of 28 haplotypes shared the most prevalent Caucasian AH8.1 (Fig. 3A), extended from the G51152/9-99431 interval into the class I region. In contrast, most of the DQB1*03-positive haplotypes, except for those carrying the DRB1*0102 allele (Fig. 3B), were broken down in the telomeric part of the class II region and had variant haplotypes in the class III (Fig. 3, D–F) and class I (data not shown) regions. Among the affected 10-3-6-8 homozygotes, only 9 of 28 haplotypes shared telomeric class III alleles and only 5 of them class I alleles (Fig. 3C). To map the telomeric border of this predisposing haplotype, we analyzed cSNP haplotypes in BTNL2 (BTII), Fig. 1) and found that the 10-3-6-8-carrying haplotypes diversified in a region of ∼40 kb between exon 5 of BTNL2 and 9-99431 located just telomeric of HLA-DRA (Fig. 1, 3d and data not shown). Thus, the haplotype analysis clearly showed that only a subset of predisposing haplotypes permitted to narrow down IGAD1 to a smaller region and was thus informative for the fine-scale mapping (Fig. 3).

Differential value of susceptibility HLA haplotypes for the fine-scale mapping

The alignment of susceptibility STR haplotypes in affected family members (Fig. 3 and supplementary online information) showed that the majority of the DQB1*02-carrying haplotypes, including the most prevalent Caucasian AH8.1 (Fig. 3A), extended from the G51152/9-99431 interval into the class I region. In contrast, most of the DQB1*03-positive haplotypes, except for those carrying the DRB1*0102 allele (Fig. 3B), were broken down in the telomeric part of the class II region and had variant haplotypes in the class III (Fig. 3, D–F) and class I (data not shown) regions. Among the affected 10-3-6-8 homozygotes, only 9 of 28 haplotypes shared telomeric class III alleles and only 5 of them class I alleles (Fig. 3C). To map the telomeric border of this predisposing haplotype, we analyzed cSNP haplotypes in BTNL2 (BTII), Fig. 1) and found that the 10-3-6-8-carrying haplotypes diversified in a region of ∼40 kb between exon 5 of BTNL2 and 9-99431 located just telomeric of HLA-DRA (Fig. 1, 3d and data not shown). Thus, the haplotype analysis clearly showed that only a subset of predisposing haplotypes permitted to narrow down IGAD1 to a smaller region and was thus informative for the fine-scale mapping (Fig. 3).

HLA-DQ/DR is the strongest predisposing locus

Although the quasi-random nature of peptide MHC class II interactions with TCR (pMHCII/TCR) may fully account for a complex inheritance of IgAD/CVID and other autoimmune disorders in families, non-MHC genes involved in cellular processes triggered by such interactions may modify the disease risk. To test this possibility and to see if the MHC is the strongest locus, we conducted an STR-based genetic linkage study with 101 multiplex families. The genome-wide scan of IgAD/CVID showed that the highest multipoint nonparametric linkage scores at 6p were not matched anywhere else in the genome (Fig. 4, Table III), implicating IGAD1 as the major disease locus. In addition to IGAD1, we found a suggestive linkage (NPL > 1.8) to non-MHC regions at 4p, 12p, 14q, 16q, 7p, and 1q (Fig. 4), and we tested these loci with a denser set of STRs and additional 109 single-case families and family-based allelic associations (Table III). Although the suggestive linkage was not supported with additional STRs at 1q and 16q (data not shown), we observed an increased sharing for adjacent STRs at 4p (NPLall = 2.45, information content 0.74, p = 0.002), 12p (NPLall = 2.23, information content 0.87, p = 0.005), and 14q (NPLall = 1.97, information content 0.81, p = 0.01; Fig. 4 and Table III). Although the p values were below the significance limit recommended for genome-wide studies, these loci were further supported by positive TDTs for STRs at 4p and 12p (Table III). Suggestive linkage at 14q and on chromosome 7 (Fig. 4), which may implicate allelic variants of the TCR genes in predisposition, is subject to further analysis in our laboratory (I. Vorechovsky, L. Hammarström, and A. D. B. Webster, manuscript in preparation).

A search for locus heterogeneity of CVID and IgAD

In HLA, we found no clear evidence for locus heterogeneity for either IgAD or CVID. Even STR haplotypes characteristic of AH8.1 were overrepresented at HLA-DQ/DR as compared with region 2 in affected family members, and we observed more family members homozygous for AH8.1 at HLA-DQ/DR than in region 2.4 In addition, CVID patients were represented among carriers of all susceptibility haplotypes (Fig. 3), supporting earlier proposals of independent investigators (5, 6, 13, 14) that the two PIDs represent a spectrum in disease severity.

Although previously reported allelic associations at IGAD1 were weaker for CVID than IgAD (7), this may have been due to a smaller number of analyzed CVID patients, occasional misdiagnosis of CVID for other PIDs (25), and locus heterogeneity, perhaps with a greater contribution from non-MHC genes in CVID than IgAD. Because rare cases of CVID-like phenotypes have been shown to result from Mendelian loss-of-function mutations (25, 41), we addressed the latter possibility by analyzing candidate genes previously implicated humoral defects for mutations (Table IV). This was conducted also for genes located in chromosome regions not exhibiting an increased allele sharing (Fig. 4), because rare sporadic cases or families would not be expected to contribute to detectable association or linkage. The mutation screening of the gene-coding regions did not identify any disease-specific changes for XBP-1, ICOSI (AILIM), OBF-1, CD38, and CTSS (Table IV), lending further support to the notion that most phenotypic variability observed in IgAD/CVID is, as for many autoimmune conditions, controlled by the interactions of class II molecules with antigenic peptides and Ag receptors on thymocytes and CD4+ T cells.

Discussion

Our results indicate that the HLA-DQ/DR locus is the major hereditary determinant of susceptibility to IgAD/CVID. The evidence is provided collectively by 1) the highest NPLZn scores at 6p21 in the genome-wide scan (Fig. 4); 2) the strongest allelic associations in the class II/III regions in family-based (Fig. 1) (21) and case-control studies (for review see Ref. 2); 3) haplotype analysis combined with the high-resolution HLA typing (Fig. 3 and Table II); 4) a higher risk of developing IgAD/CVID in HLA class II homozygotes and a significant bias in the distribution of homozgyous stretches toward the class II region (Fig. 2B); and 5) the absence of other expressed genes than HLA-DQ/DR in the interval G51152/9-99431 implicated by all fine-mapping methods (Figs. 1–3, Table I).

A number of functional lymphocyte abnormalities in vitro and substantial clinical heterogeneity suggested that CVID is a collection of Mendelian defects affecting the Ab production (42). However, neither linkage nor allelic association would have been consistently detected at IGAD1 in our patients and families (Fig. 1, Tables III–IV) and in previous case-control studies with a more limited sample size (2), had there been an extensive genetic heterogeneity of IgAD/CVID. Very few patients diagnosed as CVID have been shown to have a Mendelian defect (25), consistent with our negative mutation screening (Table IV). Furthermore, in other autoimmune disorders, the markers at 6p showed the highest excess of allelic sharing in the genome-wide scan, indicating that HLA-DQ/DR is the strongest predisposing locus (Fig. 4), with the 5-1-7-6 and 10-3-6-8 haplotypes contributing most to the disease predisposition in the analyzed population. This is consistent with
the previous observations that the frequency of susceptibility haplotype(s) in several populations is mirrored by a gradient in the population prevalence of IgAD (reviewed in Ref. 43).

In HLA, our results point to HLA-DR/DQ and not to the class III region, although not excluding a contribution of telomeric gene(s) as proposed for AH8.1 (44) and as hinted by an occasional class III homozygosity in affected individuals heterozygous for the class II alleles and by a second peak of allelic association for STRs around HLA-B (Fig. 1). However, the former observation can be due to chance and the latter finding may be fully explained by a high LD of the overrepresented AH8.1. Also, allelic associations with the disease were consistently lower for the HLA class I STRs than in the HLA-DQB1/DQA region, as was reported in previous case-control studies of the HLA class I Ags (reviewed in Ref. 2). A high LD of AH8.1 renders this haplotype much less informative for the fine-mapping studies (Fig. 3A), and differential LD of HLA haplotypes

Table III. IGAD1 confers the strongest risk of IgAD/CVID

<table>
<thead>
<tr>
<th>CHROMOSOMAL REGION</th>
<th>6p</th>
<th>4p</th>
<th>12p</th>
<th>14q</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZLR-pairs</td>
<td>4.38</td>
<td>2.95</td>
<td>2.48</td>
<td>2.27</td>
</tr>
<tr>
<td>ZLR-all</td>
<td>4.31</td>
<td>3.24</td>
<td>2.70</td>
<td>2.59</td>
</tr>
<tr>
<td>ZLR-robdom</td>
<td>3.87</td>
<td>3.43</td>
<td>2.71</td>
<td>2.89</td>
</tr>
<tr>
<td>Information content</td>
<td>0.88</td>
<td>0.79</td>
<td>0.90</td>
<td>0.80</td>
</tr>
<tr>
<td>Recombination fraction (cM from pter)</td>
<td>68</td>
<td>30</td>
<td>37</td>
<td>13</td>
</tr>
<tr>
<td>TDTs</td>
<td>D4S403, 6, 46 vs 28 (0.04)</td>
<td>D12S85, 6, 83 vs 65 (0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STR locus, allele, T vs NT (p value)</td>
<td>D4S2926, 2, 42 vs 25 (0.04)</td>
<td>D12S163, 6, 35 vs 23 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>D4S1546, 6, 68 vs 41 (0.006)</td>
<td>D12S1692, 9, 27 vs 10 (0.003)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

a ZLR scores were defined previously (30, 31).
b TDTs at 6p were reported previously (21).
well-illustrated by our results should favor gene-mapping strategies that consider all founder haplotypes, such as the trimmed haplotype method combined with homozygosity mapping (Table I, Fig. 3). The high LD of AHR8.1 may have contributed to provisional placements of the susceptibility locus to more telomeric regions (6, 16–18).

In addition to the strong genetic evidence for the major role of the HLA-DQ/DR locus, both laboratory and clinical features of CVID provide indirect support for these molecules. Because functional correlates of antigenic pMHCII/TCR interactions are distinct in mature T cells (activation, anergy) and developing thymocytes (positive and negative selection), one would expect phenotypic features reflecting both developmental stages. Indeed, a number of functional abnormalities has been observed in CVID, including restricted TCR repertoires (45, 46), increased apoptosis (47–51), enhanced sensitivity of T cells to corticosteroids (47), an increased LD of processes pertinent to autoimmunity, hypogammaglobulinemia, infection, and inflammation in symptomatic patients.

Clinical features of IgAD/CVID also resemble those found in other autoimmune disorders, including 1) the occurrence of anti-IgA Ab that are present at a higher frequency in patients with autoimmune complications than in asymptomatic IgAD (65); 2) the increased prevalence of autoimmune disease in IgAD/CVID (65); 3) frequent disease manifestation after an episode of other autoimmune phenomena; 4) variable disease severity; and 5) a slow, gradual appearance of hypogammaglobulinemia over many years (12). A transmission of IgAD to a healthy sibling by bone marrow transplantation (66) may have been due to the transfer of autoreactive T cell clones. IgAD can be transmitted from mother to fetus (7, 67), as has been observed for a number of autoimmune disorders. An autoreactive pathogenesis can also explain occasional recovery of CVID, including HIV-induced remission (68), that may have been due to the down-regulation of proinflammatory cytokines, reduction of the surface expression of HLA-DQ/DR (68), and/or altered β-chain usage of the variable portion of TCR by deleting specific Vβ families (69). The central role of HLA-DQ/DR in IgAD/CVID predisposition is also indirectly supported by the existence of humoral defects induced by drugs that interfere with Ag presentation as proposed earlier (70) and by hypogammaglobulinemia induced by the targeted inactivation of several components of Ag presentation pathways in the mouse, such as cathespin S (70), indicating the critical dependence of humoral immunity on how Ags access the immune system. A diminished number of B cells observed in a subset of CVID may be due to the autodestruction of B cells expressing IgA/IgG, but it might also develop as a result of an impaired balance between class II chain pairing (71).

Given the central predisposing role of HLA-DQ/DR, susceptibility to IgAD/CVID may be influenced by genes modifying

---

**Table IV. Mutation analysis of candidate genes for CVID**

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>MIM No.</th>
<th>Chromosome Location</th>
<th>Gene Function</th>
<th>Gene Expression</th>
<th>Mouse Phenotype[^a^]</th>
<th>CVID Patients Screened</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OBF1 (OCA-B, BOB1)</strong></td>
<td>601206</td>
<td>11q23</td>
<td>Transcriptional coactivator</td>
<td>B cells</td>
<td>Reduction of mature B cells, reduction of IgA, IgG1, IgG2a, IgG2b, but not IgM (83)</td>
<td>47 unrelated patients</td>
</tr>
<tr>
<td><strong>CD38</strong></td>
<td>107270</td>
<td>4p15</td>
<td>Glycolyase, B cell coactivator, adhesion</td>
<td>Haematopoietic cells</td>
<td>Impaired TD Ab responses (84)</td>
<td>24 patients, including families with NPL &gt; 0 for 4p STRs</td>
</tr>
<tr>
<td><strong>BST1 (CD157)</strong></td>
<td>600387</td>
<td>4p15</td>
<td>Receptor and enzymatic function (cyclase, hydrolase)</td>
<td>Haematopoietic cells</td>
<td>Reduced IgG3 in response to TI-2 Ag, reduced IgA and IgG to mucosal TD Ig, delayed B1 cells (85)</td>
<td>24 patients, including families with NPL &gt; 0 for 4p STRs</td>
</tr>
<tr>
<td><strong>ICOS (AILIM)</strong></td>
<td>604558</td>
<td>2q33</td>
<td>Costimulatory molecule, member of CD28/CTLA4 family</td>
<td>Activated T cells</td>
<td>Defective T cell activation and proliferation, reduced IgG1 and IgG2a, IL2, IL-4, and IL5, small lymph nodes and GC, susceptibility to autoimmunity (86)</td>
<td>47 unrelated patients</td>
</tr>
<tr>
<td><strong>CTSS</strong></td>
<td>116845</td>
<td>1q21</td>
<td>Papain-type cysteine protease</td>
<td>APCs</td>
<td>Impaired Ab switching to IgG2a and IgG3 (70)</td>
<td>47 unrelated patients</td>
</tr>
<tr>
<td><strong>XBP-1</strong></td>
<td>194355</td>
<td>22pter-q13</td>
<td>Leucin zipper transcription factor, binds regulatory sequences of class II genes</td>
<td>Near ubiquitous</td>
<td>Death in utero; defects in generation of plasma cells and Ig secretion (including IgM) in XBP-1-deficient lymphocytes of RAG-2 complementation system, failure to upregulate CD40L (87)</td>
<td>47 unrelated patients</td>
</tr>
<tr>
<td><strong>CD69</strong></td>
<td>107273</td>
<td>12p13</td>
<td>C-type lectin receptor</td>
<td>Activated B, T and NK cells</td>
<td>Altered pre-B cell compartment and Ig levels</td>
<td>47 patients, including families with NPL ≥ 0 at 12p STRs</td>
</tr>
</tbody>
</table>

[^a^] TD, Thymus dependent; TI, thymus independent; GC, germinal centers.
pMHCII/TCR interactions and shaping the T cell repertoire, including molecules involved in amplifying or inhibiting TCR signals, apoptosis, adhesion molecules, and coreceptors. We previously tested the aberrant splicing of *PTPRC*, a large phosphate involved in downstream TCR signaling, but no apparent association was found in CVID (72). Although variable humoral defects have been found for targeted genes involved in DNA repair (73), candidates supported by occasional chromosomal radiosensitivity in CVID (74), this phenomenon is perhaps more likely to reflect alterations of T cell subpopulations exhibiting differential radiosensitivity, such as mature and immature T cells (75). The presence of hypogammaglobulinemia in a subset of patients with the Wolf-Hirschhorn syndrome (76), a contiguous gene syndrome mapped to 4p16, suggests that a modifying gene and a gene closely linked to Wolf-Hirschhorn syndrome may be identical. In fact, several genes in the region have been implicated in humoral immune responses, including CD35 and CD157 (Table IV). A putative locus at 12p, which shows association/suggestive linkage in other autoimmune disorders (77), contains the NK gene complex, a putative locus at 12p, which shows association/suggestive linkage in several genes in the region have been implicated in humoral immunodeficiency. Genetic linkage of IgA deficiency to the major histocompatibility complex: evidence for allele segregation distortion, parent-of-origin penetrance differences and the role of anti-IgA antibodies in disease predisposition. Am. J. Hum. Genet. 64:1096.


