Am. J. Hum. Genet. 76:157–163, 2005

Report

Regression Mapping of Association between the Human Leukocyte Antigen Region and Graves Disease

Matthew J. Simmonds,[1,*] Joanna M. M. Howson,[3,*] Joanne M. Heward,[1] Heather J. Cordell,[3] Helen Foxall,[1] Jackie Carr-Smith,[1] Sarah M. Gibson,[1] Neil Walker,[3] Yaron Tomer,[4] Jayne A. Franklyn,[1] John A. Todd,[3] and Stephen C. L. Gough[1,2]

[1]Division of Medical Sciences, Institute of Biomedical Research, University of Birmingham, and [2]Birmingham Heartlands Hospital, Birmingham, United Kingdom; [3]Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, United Kingdom; and [4]Division of Endocrinology, Diabetes and Bone Diseases, Department of Medicine, Mount Sinai School of Medicine, New York

The human leukocyte antigen class II genes *DRB1*, *DQB1*, and *DQA1* are associated with Graves disease (GD), but, because of strong linkage disequilibrium within this region, the primary etiological variant(s) remains unknown. In the present study, 871 patients with GD and 621 control subjects were genotyped at the *DRB1*, *DQB1*, and *DQA1* loci. All three loci were associated with GD ($P = 1.45 \times 10^{-12}$, $P = 3.20 \times 10^{-5}$, and $P = 9.26 \times 10^{-12}$, respectively). Stepwise logistic-regression analysis showed that the association could be explained by either *DRB1* or *DQA1* but not by *DQB1*. To extend previous results, the amino acid sequence of the exon 2–encoded peptide-binding domain of *DRB1* was predicted for each subject, and, by use of logistic regression, each position was analyzed for association with GD. Of 102 amino acids, 70 were uninformative; of the remaining 32 amino acids, 13 were associated with GD (*P* values ranged from 2.20×10^{-4} to 1.2×10^{-12}). The strongest association was at position β74. This analysis is consistent with the possibility that position β74 of exon 2 of the DRB1 molecule may have a specific and central role in autoantigen presentation by DRB1 to T lymphocytes. However, we cannot yet exclude a primary role for *DQA1* or for other polymorphisms that affect DRB1 function or expression.

Graves disease (GD [MIM 275000]) is a common autoimmune disease that affects 0.5%–1% of Western populations (Tunbridge et al. 1977; Gough 2000) and results from a failure to maintain immune tolerance to thyroid antigens. Genomewide linkage scans have suggested the segregation of multiple susceptibility loci (Tomer et al. 1999; Sakai et al. 2001), but only two loci—the human leukocyte antigen (HLA) region on chromosome 6p21 and the cytotoxic T lymphocyte–associated 4 gene (*CTLA4* [MIM 123890]) on chromosome 2q33—have been confirmed (in case-control association studies) as playing primary roles in the susceptibility to GD. For *CTLA4*, disease susceptibility has been mapped to a noncoding 6.1-kb 3′ region of the gene (Ueda et al. 2003).

The HLA class II gene region encodes a number of heterodimeric receptors expressed on B lymphocytes, macrophages, dendritic cells, thymic epithelial cells, and activated T lymphocytes. Class II molecules are attractive primary candidates for etiological determinants of GD because of their involvement in both antigen presentation in the periphery and in thymic selection—namely, deletion of potentially autoreactive T cells and positive selection of a T-cell repertoire, some of which may be capable of recognizing self-epitopes and causing autoimmune damage in genetically susceptible individuals (Weetman and McGregor 1994). Reports of association of the HLA class II region with GD—in particular, the alleles *DRB1*03* (Bech et al. 1977) and *DQA1*0501* (Yanagawa et al. 1993) and the *DRB1*03-DQB1*02-DQA1*0501* haplotype—have been consistent in both case-control and family studies (Heward et al. 1998) and have produced relative risks (RRs) of 1.9–3.8 (Gough

Received June 8, 2004; accepted for publication October 19, 2004; electronically published November 22, 2004.

Address for correspondence and reprints: Dr. Stephen Gough, Department of Medicine, Division of Medical Sciences, Birmingham Heartlands Hospital, Bordesley Green East, Birmingham, B9 5SS, United Kingdom. E-mail: s.c.gough@bham.ac.uk

* These two authors contributed equally to this work.

© 2004 by The American Society of Human Genetics. All rights reserved.
0002-9297/2005/7601-0016$15.00

2000). *DRB1*08* has also been associated with GD in a small North American white population, but this finding awaits confirmation (Chen et al. 1999). In addition, the *DRB1*07* allele has been negatively associated (RR <1) with GD, which is suggestive of a protective effect against development of the disease (Chen et al. 1999). The primary etiological variants in the HLA region, however, remain unknown, largely because of the strong linkage disequilibrium (LD) between the class II alleles and other variants in the region and the fact that systematic mapping of the region is still in the early stages.

Unrelated white patients of British origin with GD were recruited from thyroid clinics within Great Britain, as described elsewhere (Heward et al. 1998). Patients were defined as having GD by the presence of biochemical hyperthyroidism together with either the presence of dysthyroid eye disease or a diffuse goitre and a significant titre of microsomal, thyroglobulin, or thyroid-stimulating-receptor autoantibodies. Ethnically matched control subjects with no personal history of autoimmune disease were bled at geographically matched sites. All control subjects had normal thyroid function and were negative for thyroid autoantibodies. All patients and subjects gave informed written consent; the local ethics committee approved the project.

DNA was prepared from whole blood by use of the Nucleon BACC II kit from Tepnel Life Sciences. The *HLA-DRB1* (MIM 142857), HLA-*DQB1* (MIM 604305) and *HLA-DQA1* (MIM 146880) alleles were amplified using the phototyping method of PCR, as published elsewhere (Bunce et al. 1995). Primers were obtained from Sigma-Genosys, and results were visualized on a 1% ethidium-bromide–stained agarose gel under UV light.

Unless otherwise stated, statistical analysis was performed within STATA, with routines written by David Clayton of the University of Cambridge. Odds ratios (ORs) with 95% CIs for both alleles and genotypes at the three loci were calculated using logistic regression (as suggested by Cordell and Clayton [2002]).

A total of 871 patients with GD and 621 control subjects were successfully genotyped at all the *DRB1, DQB1,* and *DQA1* loci. All three loci were associated with GD, regardless of whether the alleles ($P = 1.45 \times 10^{-12}$ for *DRB1,* $P = 3.20 \times 10^{-5}$ for *DQB1,* and $P = 9.26 \times 10^{-12}$ for *DQA1*) (table 1) or the genotypes ($P = 2.21 \times 10^{-5}$ for *DRB1,* $P = .01$ for *DQB1,* and $P = 2.72 \times 10^{-8}$ for *DQA1*) (table A1 [online only]) were modeled. The *DRB1*03* (OR 2.98; 95% CI 2.26–3.92) and *DQA1*0501–02* (OR 2.53; 95% CI 1.95–3.27) alleles were the most positively associated (with use of *DRB1*07* and *DQA1*0201* alleles, respectively, as a reference). There was some evidence of increased risk associated with the *DRB1*08* alleles, ranked second (with an OR significantly >1) to the *DRB1*03* alleles (OR 2.16; 95% CI 1.25–3.74) (table 1), with use of *DRB1*07* as a reference. An increase in the frequency of the *DRB1*03-DQB1*02-DQA1*0501* haplotype was observed in patients with GD, compared with control subjects (OR 1.45; 95% CI 1.20–1.75), whereas a decrease in the frequency of the *DRB1*07-DQB1*02-DQA1*0201* haplotype was observed in patients with GD, compared with control subjects (OR 0.51; 95% CI 0.40–0.66).

A stepwise logistic-regression approach was used to test whether the associations at the *DRB1, DQB1,* and *DQA1* loci could be explained by a single locus (Cordell

Table 1

ORs and 95% CIs for the Alleles of *HLA-DRB1, HLA-DQB1,* and *HLA-DQA1*

Gene and Allele(s)	OR (95% CI)
HLA-DRB1[a]:	
0301–5	2.98 (2.26–3.92)
1001	2.44 (.58–10.17)
0801–11	2.16 (1.25–3.74)
1601–6	2.09 (.84–5.20)
1101–21	2.03 (1.43–2.87)
0101–4	1.80 (1.32–2.45)
0401–22	1.58 (1.20–2.09)
1501–5	1.52 (1.14–2.04)
1301–22	1.45 (1.05–2.00)
1201–3	1.17 (.58–2.36)
1401–21	1.04 (.62–1.76)
0901	1.01 (.52–1.94)
0701	1.00[b]
HLA-DQB1[c]:	
02	2.56 (1.67–3.91)
04	2.88 (1.47–5.64)
0301/4	2.29 (1.49–3.52)
05	2.17 (1.39–3.38)
0601–9	1.80 (1.18–2.75)
0302	1.68 (1.05–2.68)
03032	1.00[b]
HLA-DQA1[d]:	
0501–02	2.53 (1.95–3.27)
0601	2.77 (.48–15.93)
0401	2.48 (1.37–4.47)
0101/4	1.64 (1.22–2.21)
0102–3	1.51 (1.16–1.97)
03011–12	1.49 (1.13–1.96)
0201	1.00[b]

NOTE.—Alleles *DQB1*0305* and *DQA1*0302* were detected only in one individual and so were not used in the statistical analysis.

[a] The overall *P* value for association with GD was 1.45×10^{-12}.

[b] The least susceptible allele is used as reference.

[c] The overall *P* value for association with GD was 3.20×10^{-5}.

[d] The overall *P* value for association with GD was 9.26×10^{-12}.

Table 2

Genotype and Allele ORs and 95% CIs for Protective and Susceptible Alleles at *HLA-DRB1*

DRB1 Genotype	Case: Control Subject Frequency	OR (95% CI) Using Reference		
		03/03	07/07	X/X
03	475:206	2.98 (2.26–3.92)	1.00[a]	1.85 (1.52–2.26)
X	1,122:849	1.61 (1.28–2.03)	.54 (.44–.66)	1.00[a]
07	145:187	1.00[a]	.34 (.26–.44)	.62 (.49–.78)
03/03	45:11	1.00[a]	9.09 (3.26–25.37)	3.49 (1.77–6.87)
03/X	351:154	.56 (.28–1.11)	5.06 (2.26–11.38)	1.94 (1.52–2.48)
X/X	339:289	.29 (.15–.56)	2.61 (1.17–5.81)	1.00[a]
07/03	34:30	.28 (.12–.63)	2.52 (1.00–6.37)	.97 (.58–1.62)
07/X	93:117	.19 (.10–.40)	1.77 (.77–4.06)	.68 (.49–.93)
07/07	9:20	.11 (.04–.31)	1.00[a]	.38 (.17–.86)

Note.—Category X includes all alleles not listed. *DRB1*03* codes for *DRB1*0301–28; DRB1*07* codes for *DRB1*0701–8*.

[a] These alleles were used as reference.

and Clayton 2002). Two analyses of the data were undertaken: one in which the loci were modeled using alleles and the other in which they were modeled using genotypes. The model used did not affect the outcome of the study. The most associated locus—the one with the smallest *P* value—was put in the logistic model. The other two loci were added one at a time, and a likelihood-ratio test was used to test whether either of them improved the model. A *P* value <.01 was considered suggestive of an improvement in the model. When an allele-based model containing *DRB1* was used, addition of neither *DQB1* ($P = .15$) nor *DQA1* ($P = .90$) significantly improved the model. Similarly, when *DQA1* was used in the model as the best locus, neither *DQB1* nor *DRB1* improved the model ($P = .04$ and $P = .05$, respectively). Conversely, *DQB1* was improved by both *DRB1* ($P = 1.88 \times 10^{-8}$) and *DQA1* ($P = 3.99 \times 10^{-8}$). These results indicate that *DQB1* is not a primary locus in the class II region. Similar results were obtained when the loci were modeled by genotype. With *DRB1* as the best locus, neither *DQA1* nor *DQB1* improved the model ($P = .28$ and $P = .13$, respectively). Likewise, when *DQA1* was entered into the regression model as the best locus, addition of *DRB1* or *DQB1* did not improve the model ($P = .35$ and $P = .36$, respectively). Again, *DQB1* was improved by both *DQA1* ($P = 4.06 \times 10^{-6}$) and *DRB1* ($P = 2.71 \times 10^{-5}$).

After we found that either *DRB1* or *DQA1* could explain the associations reported in table 1, we tested whether phase was important. Two-locus haplotypes were reconstructed using Sinnwell and Schaid's haplo.stats package for R, and a likelihood-ratio test was used to test for haplotype-dependent or *cis* interactions between phased and unphased genotypes. *DRB1-DQA1* haplotypes with frequency <0.5% were grouped, as were genotypes with frequency <0.5% at *DRB1*. No evidence of haplotype-specific effects was found ($P = .10$).

Allele and genotype ORs at the *DRB1* loci are shown in table 2. *DRB1*X* or *DRB1*X/DRB1*X* (where X = any *DRB1* locus except *DRB1*03* or *DRB1*07*) was used as reference. One copy of the *DRB1*03* allele increased the risk of GD (OR 1.94; 95% CI 1.52–2.48), and two copies further increased this risk (OR 3.49; 95% CI 1.77–6.87). Conversely, one copy of *DRB1*07* decreased disease risk (OR 0.68; 95% CI 0.49–0.93), and two copies further decreased this risk (OR 0.38; 95% CI 0.17–0.86). The protective effect of the *DR7* haplotype was seen even in the absence of the *DRB1*03* allele (OR 0.66; 95% CI 0.51–0.85). These data suggest that the susceptible and protective alleles are working in a multiplicative manner to increase or decrease disease risk, respectively. The results also showed that the presence of one copy of each allele (*DRB1*03/DRB1*07* heterozygote) neither increased nor decreased disease risk (OR 0.97; 95% CI 0.58–1.62), when compared with the reference genotype.

Allele and genotype ORs at the *DQA1* loci are shown in table 3. Testing *DQA1* by use of *DQA1*X* or *DQA1*X/DQA1*X* (where *X* = any *DQA1* locus except *DQA1*05* or *DQA1*02*) as reference, with an OR of 1, we saw a similar effect. A single copy of *DQA1*05* increases the risk of GD (OR 2.10; 95% CI 1.63–2.70), whereas two copies further increase the risk (OR 2.31; 95% CI 1.55–3.42). In contrast to *DRB1*07,* two copies of *DQA1*02* are needed to significantly decrease the risk of GD (OR 0.34; 95% CI 0.14–0.82). The results also showed that the presence of one copy of each allele (*DQA1*05/DQA1*02* heterozygote) neither increased nor decreased disease risk (OR 1.00; 95% CI 0.64–1.55), when compared with the reference ge-

Table 3

Genotype and Allele ORs and 95% CIs for Protective and Susceptible Alleles at *HLA-DQA1*

DQA1 Genotype	Case: Control Subject Frequency	OR (95% CI) Using Reference		
		05/05	02/02	X/X
05	645:316	1.00[a]	2.52 (1.95–3.27)	1.63 (1.37–1.93)
X	954:746	.61 (.52–.73)	1.55 (1.22–1.96)	1.00[a]
02	143:180	.40 (.31–.51)	1.00[a]	.65 (.51–.82)
05/05	104:44	1.00[a]	6.75 (2.66–17.12)	2.31 (1.55–3.42)
05/X	389:181	.91 (.61–1.35)	6.14 (2.55–14.78)	2.10 (1.63–2.70)
X/X	242:236	.43 (.29–.64)	2.93 (1.22–7.06)	1.00[a]
05/02	48:47	.43 (.25–.74)	2.92 (1.13–7.55)	1.00 (.64–1.55)
02/X	81:93	.37 (.23–.58)	2.49 (1.00–6.19)	.85 (.60–1.20)
02/02	7:20	.15 (.06–.38)	1.00[a]	.34 (.14–.82)

Note.—Category X includes all alleles not listed. *DQA1*05* codes for *DQA1*0501–5; DQA1*02* codes for *DQA1*0201.*

[a] These alleles were used as reference.

notype. The presence of the *DRB1*07* haplotype, therefore, appears to cancel out *DRB1*03* susceptibility. Several mechanisms are possible, such as (1) epitope stealing by *DRB1*07* haplotype DR or DQ molecules, (2) dominant negative selection in the thymus of a *DRB1*03* autoantigenic epitope, (3) *DRB1*07* haplotype–driven selection of protective T-regulatory cells in the thymus, or (4) some other unique feature of the *DRB1*07* haplotype, *DRB1,* or *DQA1* alleles, such as regulation of their expression.

To extend previous results that specifically implicate *DRB1* position β74 in GD susceptibility (Ban et al. 2004), we used the regression methods in our larger data set to predict the specific amino acid identity of the exon 2–encoded domain of DRB1 for each subject. We used amino acid sequences of the mature protein of the different DRB1 molecules from the IMGT/HLA Sequence Database. Since exon 2 varied in size between the different *DRB1* subtypes, all subtypes were aligned to *DRB1*0301,* such that position 1 of each *DRB1* type was relative to the position of the first amino acid in exon 2 that formed the DRB1 peptide–binding environment of *DRB1*0301.* When all the *DRB1* exons were aligned, a 102-aa sequence map was produced. Because of the resolution limitations of the phototyping method used to analyze the *DRB1* locus (Bunce et al. 1995), a generic sequence of the *DRB1* molecule was constructed for the given loci with amino acid sequence variation in a group of specific subtypes, and the amino acid at that given position was coded as "missing." For example, our subtyping method could not distinguish between the different *DRB1*04* subtypes; therefore, at amino acid position β74, where there is the potential to have glutamate (Glu), alanine (Ala), leucine (Leu), or arginine (Arg) (for *DRB1*0406, DRB1*0408, DRB1*0412,* and *DRB1*0422,* respectively), the amino acid was scored as a zero at this position in all *DRB1*04*-positive subjects. Likewise, at position β74, we were unable to distinguish Leu from Ala, the amino acids corresponding to *DRB1*08* alleles. We also established a genotyping assay to determine Arg positivity at position β74. Since Arg is present only in individuals with *DRB1*03* (except *DRB1*0311, DRB1*0317,* and *DRB1*0324*) and with subtypes *DRB1*0422* and *DRB1*1107,* these results enabled us to infer results for a larger proportion of our data set, dependent on *DRB1* status obtained from previous typing. Amplicons that contained the *DRB1*-binding domain β74 were amplified with forward primer 5′-TAC CGG GCG GTG ACG GAG TG-3′ and reverse primer 5′-CAC CAA CCC CGT AGT TGT GTC TGC-3′ (Sigma-Genosys) within an MJ Research thermal cycler, by use of a program that consisted of an initial denaturation step of 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 30 s. A final step of 72°C for 5 min was performed. The reactions were performed in a volume of 25 μl that contained 200 ng of DNA, 25 pmol of each primer, 0.3 U *Taq* polymerase, 2 mM $MgCl_2$, and 200 μM dNTPs. The amplicon was 121 bp in length and was visualized on a 2% agarose gel and then was subjected to RFLP with *Msp*I, in accordance with manufacturer conditions (New England Biolabs). The products were then visualized on a 3% agarose gel, and unambiguous genotypes were assigned accordingly.

Logistic-regression analysis was used to perform a single-locus association analysis at each of the amino acid positions encoded by *DRB1* exon 2. Amino acids were modeled as alleles, which assumed a multiplicative mode of inheritance, and also as genotypes, in which no particular mode of inheritance was assumed. Since the multiplicative model relied on fewer degrees of freedom (the number of alleles minus 1) compared with the genotype

model, which requires more degrees of freedom (the number of genotypes minus 1), the multiplicative model is more powerful and therefore was preferred. However, a likelihood-ratio test was performed to verify that there was no significant difference at the 5% level between the two models. When there was a significant difference, the full model was employed. Of the 102 amino acid positions analyzed, 66 were nonpolymorphic. A further four amino acids positions—β51, β57, β60, and β77—had three or fewer polymorphisms in the control populations and had no polymorphisms in the cases. For the purpose of this analysis, therefore, these four amino acids were also classed as nonpolymorphic, so our data set comprised a total of 70 nonpolymorphic or uninformative amino acids. Logistic-regression disease-association analysis was performed for the remaining 32 amino acid positions. Because of the large number of tests being performed, a P value of .001 was used as a cutoff for evidence of association. With that criterion, the regression analysis identified 13 amino acid positions associated with disease (β10, β11, β12, β13, β14, β25, β26, β30, β32, β47, β58, β71, and β74), with P values ranging from 2.20×10^{-4} for β58 to 1.22×10^{-12} for β74.

A two-locus stepwise logistic-regression analysis (Cordell and Clayton 2002) was performed for a number of the polymorphic amino acid positions, in an attempt to evaluate the effects of LD between the various alleles that encode these nonsynonymous SNPs. Because of the missing data at certain positions, we confined our analysis to loci with complete information in at least 300 informative cases and 300 control subjects, which resulted in 98% power, with an allele frequency of 10%, an OR of 2, and a P value of .05. Eleven amino acid positions (β12, β14, β25, β27, β31, β32, β33, β62, β74, β76, and β82), including five of the disease-associated amino acids, fulfilled this criterion and were appropriate for the two-locus analysis. Since β74 was the most associated amino acid ($P = 1.2 \times 10^{-12}$), we selected that position as the best locus; we employed stepwise logistic regression to test whether any amino acid positions improved the model when β74 was included. Each amino acid position was added in turn; none of the 10 remaining amino acids improved the model ($P = .05$–.64). Stepwise logistic-regression analysis was employed again, to test whether β74 improved a model that included each of the other 10 amino acid positions. Each model was tested in turn, and position β74 improved every other locus ($P = .004$ to $P = 5.4 \times 10^{-13}$), which indicates that LD with β74 can account for the disease associations of the other 10 amino acid positions of *DRB1*.

Our data, therefore, are consistent with the possibility that *DRB1* position β74 is a primary determinant of susceptibility to GD. However, even though we have used both a larger sample size than those of previous studies and efficient regression approaches, we cannot rule out the possibility that the HLA association is due to variation at the *DQA1* locus, to a combination of *DQA1* and *DRB1* variation, or to some other variation at the *DRB1* locus, including other SNPs within exon 2 or variation outside exon 2. Moreover, at position β74, no other *DRB1* molecules except *DRB1*03* (and subtypes *DRB1*0422* and *DRB1*1107*) contain a positively charged Arg at β74, and only *DRB1*07* allotypes (and *DRB1*0311, DRB1*0317,* and *DRB1*0324* subtypes of *DRB1*03*) contain a noncharged polar Gln at that position. Unfortunately, alleles that could help further test the candidacy of β74, such as *DRB1*0422* and *DRB1*1107,* were too rare in our sample (0.7% of cases and 0.4% of controls) to allow meaningful statistical analysis. The inability to fully resolve phase-known allele status remains a significant limitation in these fine-mapping studies. *DRB1*07* also contains amino acids at positions β11, β13, and β30 (glycine [Gly], phenylalanine [Phe], and Leu, respectively) that are different from the rest of the *DRB1* types, but these do not vary significantly in either size or charge, in comparison with the other nonassociated *DRB1* subtypes. The positive association of *DRB1*08* alleles (Leu or Ala at position β74) suggests that the simple model of Arg versus Gln at position β74 cannot fully explain the association of the HLA class II region with GD. Nevertheless, our finding that the phase of *DRB1-DQA1* haplotypes is not a significant factor in the HLA association with GD in the sample studied does reduce the likeliness of the existence of an additional disease locus that was untyped in our study and is in LD with *DRB1* or *DQA1*.

Notwithstanding the caveats discussed above, findings of prior genetic and functional studies make β74 a strong candidate. It has been suggested that β74 directly influences TCR binding, which provides a site-specific mimic of a side-chain recognition site sufficient for TCR activation (Nepom et al. 1996). Studies of CLIP (which is a cleaved peptide from the invariant chain that participates in class II peptide-binding functions) accumulation in binding grooves of *DRB1* molecules have shown that when the Arg that is present at β74 is mutated to Gln, the binding ability of the molecule to CLIP is increased (Doebele et al. 2003). Amino acid position β74 also spans several binding environments involved in autoantigen binding/presentation and T-cell antigen-receptor docking and interaction (Chelvanayagam 1997). Regarding type 1 diabetes (T1D [MIM 222100]), it has been established that *DRB1*0403* and *DRB1*0406* alleles provide a lower risk of disease, compared with the high-risk *DRB1*0401* allele (Awata et al. 1992; Cucca et al. 1993, 1995, 2001), with variation at position β74 in the P4-binding pocket seen between the lower-risk *DRB1*0403* and *DRB1*0406* alleles, which contain a negatively charged Glu, compared with the high-risk

*DRB1*0401* allele, which contains a noncharged polar Ala. Position β74 also encompasses part of the shared epitope region in the P4-binding pocket that is associated with rheumatoid arthritis (RA [MIM 180300]) (Gregersen et al. 1987). The *DRB1*07* haplotype is also negatively associated with T1D and RA (Cavan et al. 1993; Weyand et al. 1995; Zhao et al. 1996). Complete sequence determination in very large sample sizes and additional functional studies will be required to further distinguish between the various explanations for the association of the HLA region with GD and other HLA-associated disorders.

Acknowledgments

We acknowledge the Wellcome Trust, the Juvenile Diabetes Research Foundation, and the special trustees of the former United Birmingham Hospitals for their support. We thank all thyroid-research nurses, for assistance with data collection, and consultants at U.K. thyroid clinics, for access to their patients.

Electronic-Database Information

The URLs for data presented herein are as follows:

David Clayton's routines for STATA, http://www-gene.cimr.cam.ac.uk/clayton/software/stata/

haplo.stats package, http://cran.r-project.org/src/contrib/Descriptions/haplo.stats.html (for v1.1.1 for R v1.8.1)

IMGT/HLA Sequence Database, http://www.ebi.ac.uk/imgt/hla/links.html

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for GD, *CTLA4, HLA-DRB1, HLA-DQB1, HLA-DQA1,* T1D, and RA)

STATA 8, http://www.stata.com/

References

Awata T, Kuzuya T, Matsuda A, Iwamoto Y, Kanazawa Y (1992) Genetic analysis of HLA class II alleles and susceptibility to type 1 (insulin-dependent) diabetes mellitus in Japanese subjects. Diabetologia 35:419–424

Ban Y, Davies TF, Greenberg DA, Concepcion ES, Osman R, Oashi T, Tomer Y (2004) Arginine at position 74 of the HLA-DR β1 chain is associated with Graves' disease. Genes Immun 5:203–208

Bech K, Lumholtz B, Nerup J, Thomsen M, Platz P, Ryder LP, Svejgaard A, Siersbaek-Nielsen K, Hansen JM, Larsen JH (1977) HLA antigens in Graves' disease. Acta Endocrinol (Copenh) 86:510–516

Bunce M, O'Neill CM, Barnardo MC, Krausa P, Browning MJ, Morris PJ, Welsh KI (1995) hototyping: comprehensive DNA typing for *HLA-A, B, C, DRB1, DRB3, DRB4, DRB5 & DQB1* by PCR with 144 primer mixes utilizing sequence-specific primers (PCR-SSP). Tissue Antigens 46:355–367

Cavan DA, Jacobs KH, Penny MA, Kelly MA, Mijovic C, Jenkins D, Fletcher JA, Barnett AH (1993) Both DQA1 and DQB1 genes are implicated in HLA-associated protection from type 1 (insulin-dependent) diabetes mellitus in a British Caucasian population. Diabetologia 36:252–257

Chelvanayagam G (1997) A roadmap for HLA-DR peptide binding specificities. Hum Immunol 58:61–69

Chen QY, Huang W, She JX, Baxter F, Volpe R, Maclaren NK (1999) HLA-DRB1*08, DRB1*03/DRB3*0101, and DRB3*0202 are susceptibility genes for Graves' disease in North American Caucasians, whereas DRB1*07 is protective. J Clin Endocrinol Metab 84:3182–3186

Cordell HJ, Clayton DG (2002) A unified stepwise regression procedure for evaluating the relative effects of polymorphisms within a gene using case/control or family data: application to *HLA* in type 1 diabetes. Am J Hum Genet 70:124–141

Cucca F, Lampis R, Congia M, Angius E, Nutland S, Bain SC, Barnett AH, Todd JA (2001) A correlation between the relative predisposition of MHC class II alleles to type 1 diabetes and the structure of their proteins. Hum Mol Genet 10:2025–2037

Cucca F, Lampis R, Frau F, Macis D, Angius E, Masile P, Chessa M, Frongia P, Silvetti M, Cao A, DeVirgilis S, Congia M (1995) The distribution of DR4 haplotypes in Sardinia suggests a primary association of type I diabetes with DRB1 and DQB1 loci. Hum Immunol 43:301–308

Cucca F, Muntoni F, Lampis R, Frau F, Argiolas L, Silvetti M, Angius E, Cao A, De Virgiliis S, Congia M (1993) Combinations of specific DRB1, DQA1, DQB1 haplotypes are associated with insulin-dependent diabetes mellitus in Sardinia. Hum Immunol 37:85–94

Doebele RC, Pashine A, Liu W, Zaller DM, Belmares M, Busch R, Mellins ED (2003) Point mutations in or near the antigen-binding groove of HLA-DR3 implicate class II-associated invariant chain peptide affinity as a constraint on MHC class II polymorphism. J Immunol 170:4683–4692

Gough SC (2000) The genetics of Graves' disease. Endocrinol Metab Clin North Am 29:255–266

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

Heward JM, Allahabadia A, Daykin J, Carr-Smith J, Daly A, Armitage M, Dodson PM, Sheppard MC, Barnett AH, Franklyn JA, Gough SC (1998) Linkage disequilibrium between the human leukocyte antigen class II region of the major histocompatibility complex and Graves' disease: replication using a population case control and family-based study. J Clin Endocrinol Metab 83:3394–3397

Nepom GT, Ou D, Lybrand TP, DeWeese C, Domeier ME, Buckner JH, Mitchell LA, Tingle AJ (1996) Recognition of altered self major histocompatibility complex molecules modulated by specific peptide interactions. Eur J Immunol 26:949–952

Sakai K, Shirasawa S, Ishikawa N, Ito K, Tamai H, Kuma K, Akamizu T, Tanimura M, Furugaki K, Yamamoto K, Sasazuki T (2001) Identification of susceptibility loci for autoimmune thyroid disease to 5q31-q33 and Hashimoto's thyroiditis to 8q23-q24 by multipoint affected sib-pair linkage analysis in Japanese. Hum Mol Genet 10:1379–1386

Tomer Y, Barbesino G, Greenberg DA, Concepcion E, Davies

TF (1999) Mapping the major susceptibility loci for familial Graves' and Hashimoto's diseases: evidence for genetic heterogeneity and gene interactions. J Clin Endocrinol Metab 84:4656–4664

Tunbridge WM, Evered DC, Hall R, Appleton D, Brewis M, Clark F, Evans JG, Young E, Bird T, Smith PA (1977) The spectrum of thyroid disease in a community: the Whickham survey. Clin Endocrinol (Oxf) 7:481–493

Ueda H, Howson JM, Esposito L, Heward J, Snook H, Chamberlain G, Rainbow DB, et al (2003) Association of the T-cell regulatory gene *CTLA4* with susceptibility to autoimmune disease. Nature 423:506–511

Weetman AP, McGregor AM (1994) Autoimmune thyroid disease: further developments in our understanding. Endocr Rev 15:788–830

Weyand CM, McCarthy TG, Goronzy JJ (1995) Correlation between disease phenotype and genetic heterogeneity in rheumatoid arthritis. J Clin Invest 95:2120–2126

Yanagawa T, Mangklabruks A, Chang YB, Okamoto Y, Fisfalen ME, Curran PG, DeGroot LJ (1993) Human histocompatibility leukocyte antigen-DQA1*0501 allele associated with genetic susceptibility to Graves' disease in a Caucasian population. J Clin Endocrinol Metab 76:1569–1574

Zhao Y, Dong Y, Zhu X, Qiu C (1996) HLA-DRB1 alleles genotyping in patients with rheumatoid arthritis in Chinese. Chin Med Sci J 11:232–235