Next Generation Sequencing Reveals the Association of DRB3*02:02 With Type 1 Diabetes

Henry A. Erlich,[1] Ana Maria Valdes,[2] Shana L. McDevitt,[3] Birgitte B. Simen,[4] Lisbeth A. Blake,[4] Kim R. McGowan,[4] John A. Todd,[5] Stephen S. Rich,[6] and Janelle A. Noble,[3] for the Type 1 Diabetes Genetics Consortium (T1DGC)*

The primary associations of the HLA class II genes, HLA-DRB1 and HLA-DQB1, and the class I genes, HLA-A and HLA-B, with type 1 diabetes (T1D) are well established. However, the role of polymorphism at the HLA-DRB3, HLA-DRB4, and HLA-DRB5 loci remains unclear. In two separate studies, one of 500 subjects and 500 control subjects and one of 366 DRB1*03:01–positive samples from selected multiplex T1D families, we used Roche 454 sequencing with Conexio Genomics ASSIGN ATF 454 HLA genotyping software analysis to analyze sequence variation at these three HLA-DRB loci. Association analyses were performed on the two HLA-DRB loci haplotypes (DRB1-DRB3, -DRB4, or -DRB5). Three common HLA-DRB3 alleles (*01:01, *02:02, *03:01) were observed. DRB1*03:01 haplotypes carrying DRB3*02:02 conferred a higher T1D risk than did DRB1*03:01 haplotypes carrying DRB3*01:01 in DRB1*03:01/*03:01 homozygotes with two DRB3*01:01 alleles (odds ratio [OR] 3.4 [95% CI 1.46–8.09]), compared with those carrying one or two DRB3*02:02 alleles (OR 25.5 [3.43–189.2]) (P = 0.033). For DRB1*03:01/*04:01 heterozygotes, however, the HLA-DRB3 allele did not significantly modify the T1D risk of the DRB1*03:01 haplotype (OR 7.7 for *02:02; 6.8 for *01:01). These observations were confirmed by sequence analysis of HLA-DRB3 exon 2 in a targeted replication study of 281 informative T1D family members and 86 affected family-based association control (AFBAC) haplotypes. The frequency of DRB3*02:02 was 42.9% in the DRB1*03:01/*03:01 patients and 27.6% in the DRB1*03:01/*04 (P = 0.005) compared with 22.6% in AFBAC DRB1*03:01 chromosomes (P = 0.001). Analysis of T1D-associated alleles at other HLA loci (HLA-A, HLA-B, and HLA-DPB1) on DRB1*03:01 haplotypes suggests that DRB3*02:02 on the DRB1*03:01 haplotype can contribute to T1D risk. **Diabetes 62:2618–2622, 2013**

From [1]Roche Molecular Systems, Pleasanton, California; [2]King's College London, London, U.K.; the [3]Children's Hospital Oakland Research Institute, Oakland, California; [4]454 Life Sciences–a Roche Company, Branford, Connecticut; the [5]Department of Medical Genetics, Cambridge Institute of Medical Research, JDRF/Wellcome Trust Diabetes and Inflammation Laboratory, National Institute for Health Research Cambridge Biomedical Research Centre, University of Cambridge, Addenbrooke's Hospital, Cambridge, U.K.; and the [6]Center for Public Health Genomics, University of Virginia, Charlottesville, Virginia.
Corresponding author: Henry A. Erlich, henry.erlich@roche.com.

The Type 1 Diabetes Genetics Consortium (T1DGC) (1) is an international collaboration that has collected thousands of multiplex family samples, as well as case and control samples, and has carried out linkage and association analysis for genome-wide single nucleotide polymorphisms (SNPs), candidate gene SNPs, and major histocompatibility complex (MHC) SNPs, as well as for alleles and haplotypes at the highly polymorphic HLA class I and class II genes (2–7). More than 40 different loci have been associated with T1D (3); however, linkage and association analyses have identified the HLA region as the major genetic determinant of T1D risk. The most strongly T1D-associated MHC markers are defined by the HLA-DRB1, -DQA1, and -DQB1 haplotypes (4), although alleles at other HLA loci, notably HLA-A and HLA-B, as well as HLA DPB1 (5–11) and other non-HLA loci across the genome, also contribute to T1D risk (3).

Although the association of HLA-DRB1 alleles with T1D is well established, the role of polymorphism at the HLA-DRB3, HLA-DRB4, and HLA-DRB5 loci has been studied less frequently, partly due to technical issues in genotyping. Most high-resolution genotyping strategies for HLA-DRB1 depend on DRB1-specific PCR primers to minimize confounding signals from coamplified secondary DRB loci. All copies of chromosome 6 have a DRB1 locus, and most, but not all, have a functional second DRB locus. This second DRB locus is DRB3 for DRB1*03, *11, *12, *13, and *14 haplotypes, DRB4 for DRB1*04, *07, and *09 haplotypes, and DRB5 for DRB1*15 and *16 haplotypes. DRB1*01, DRB1*08, and DRB1*10 haplotypes do not have a second DRB locus.

The clonal sequencing property of next generation sequencing systems, such as the Roche 454 GS FLX and GS Junior Systems, allows the use of generic DRB primers to coamplify and sequence exon 2 from DRB1, DRB3, DRB4, and DRB5 loci (12–14). We have used the Roche 454 amplicon sequencing system with "fusion primers" containing the 454 adaptor (A and B) sequences and 10 base multiplex ID tags (MIDs) to amplify and sequence exon 2 from different individuals (13,14). The genotyping software consolidates these clonal sequence readings, sorts them to individual DRB loci and to individual samples, and compares them to the IMGT/HLA Sequence Database to assign specific genotypes at the HLA-DRB1, HLA-DRB3, HLA-DRB4, and HLA-DRB5 loci.

To assess the potential role of these secondary HLA-DRB loci in T1D, association analyses were carried out at the two DRB locus haplotype levels, focusing on the DRB1*03:01 haplotypes bearing DRB3*01:01 or DRB3*02:02. The role of HLA-DRB3 polymorphism in T1D risk was initially evaluated in a case/control study and then examined in a targeted replication study of informative patients from

multiplex HLA-genotyped families, allowing analysis of HLA haplotypes and the potential role of T1D-associated alleles at other HLA loci.

RESEARCH DESIGN AND METHODS

Subjects. The initial set of DNA samples from 500 case and 500 control subjects was provided by the JDRF/Wellcome Trust Diabetes and Inflammation Laboratory (case subjects) and the British 1958 Birth Cohort (control subjects) from a study described previously (15).

To test the hypothesis generated in the case/control study, DNA samples from the T1DGC and the Human Biological Data Interchange (HBDI) multiplex family collections were used. Samples from 280 T1D-affected family members were selected if they had a DRB1*03 haplotype. Samples were selected based on the known T1D risk DRB1 genotypes, DRB1*03:01/DRB1*03:01 (n = 54 T1D subjects; 105 independent DRB1*03:01 haplotypes), DRB1*03:01/DRB1*04:01 (n = 154 DRB1*03:01 haplotypes), and DRB1*03:01/DRB1*04:04 (n = 73 DRB1*03:01 haplotypes), with the aim of testing the hypothesis that DRB3*02:02 and *01:01 alleles may confer differential T1D risk. Parent samples with a DRB1*03:01 allele known not to be transmitted to an affected T1D patient were also selected as control subjects from the T1DGC and the HBDI collections (affected family-based association controls [AFBACs], n = 115 haplotypes) (16). These AFBAC control haplotypes were used to determine population frequencies of DRB3*02:02 and DRB3*01:01 alleles on DRB1*03:01 haplotypes.
HLA genotyping using next generation sequencing (Roche 454). HLA sequence data were generated on the Roche 454 GS FLX and GS Junior Systems and analyzed using Conexio Genomics HLA ASSIGN ATF genotyping software to interpret the sequence files as HLA genotypes (13,14). Amplicons were generated from genomic DNA using DRB generic exon 2 454 fusion primers. The 454 fusion primers consist of a locus-specific sequence on the 3′ end, a 10-bp MID tag, and an "A" or "B" 454-specific primer sequence on the 5′ end. The MID tag serves as a sample barcode recognized by Conexio ASSIGN ATF genotyping software.

Amplicons were purified with AMPure beads (Becton Dickinson, Franklin Lakes, NJ), quantified using the Quant-iT PicoGreen dsDNA reagent (Life Technologies, Foster City, CA), and mixed with capture beads after dilution. Individual DRB exon 2 amplicon molecules captured by these beads were amplified in an emulsion PCR amplification and DNA-containing beads subsequently analyzed by pyrosequencing to obtain sequence readings originating from a single molecule (12–14). Sequencing on the GS FLX and GS Junior System was performed, as described (13,14). HLA genotypes were assigned to samples using Conexio ASSIGN ATF genotyping software, as described (13,14).
Statistical analysis. DRB3 allele frequencies on DR3 haplotypes in DR3 homozygotes, DR3/DR4 heterozygotes, and AFBACs were compared using a Fisher exact test where the overall sample size was less than 50 or a Pearson χ^2 statistic.

RESULTS

Case/control study. A total of 500 case and 500 control DNA samples from the Wellcome Trust/JDRF Diabetes Inflammation Laboratory collection were amplified with 454 DRB fusion primers containing 32 MID tags and sequenced in two GS FLX System runs using PicoTiterPlates fitted with 16-region gaskets. The long read lengths of >400 bp spanned the amplicon in both directions and allowed setting phase (haplotyping) for the polymorphisms within exon 2. This provided, in most cases, unambiguous genotype assignments for *HLA-DRB1*, *HLA-DRB3*, and *HLA-DRB5*. For *HLA-DRB4*, however, several different genotypes were consistent with the observed sequence reads (ambiguity string). (The *HLA-DRB4* exon 2 sequence reads were identical for the DRB4*01:01, DRB4*01:03, and DRB4*01:06 alleles.) This DRB4 exon 2 sequence was present on all DRB1*04, *07, and *08 haplotypes; thus, the role of DRB4 alleles for association with T1D or for defining specific linkage disequilibrium (LD) patterns could not be evaluated.

For *HLA-DRB3* and *HLA-DRB5*, the allele assignments were unambiguous and, as expected, a pattern of very strong LD between *HLA-DRB1* and the secondary DRB locus was observed. The T1D association analyses were performed on the two-locus haplotypes (Table 1). In these data, DRB1*15:01 was linked exclusively to DRB5*01:01,

and although the numbers were small (n = 6), DRB1*16:01 was linked exclusively to DRB5*02:02. Three common DRB3 alleles (*01:01, *02:02, *03:01) were observed. Of those haplotypes carrying a DRB3 locus, most DRB1 alleles were linked to a unique DRB3 allele in both case and control subjects (Table 1). The DRB1*03:01 and the DRB1*13:01 haplotypes, however, were linked to one of two different DRB3 alleles (*01:01 or *02:02).

Thus, the role of DRB3*01:01 versus DRB3*02:02 on DRB1*03:01 haplotypes could be evaluated. DRB1*03:01 haplotypes carrying DRB3*02:02 had a nominally higher risk for T1D; this difference in T1D risk was observed in the DRB1*03:01/*03:01 homozygotes with two DRB3*01:01 alleles (odds ratio [OR] 3.4 [95% CI 1.46–8.09]) compared with those with one or two DRB3*02:02 alleles (25.5 [3.43–189.2]; P = 0.033; Table 2). For DRB1*03:01/*04:01 heterozygotes, however, there was no difference in the T1D risk between DRB1*03:01 haplotypes distinguished by the DRB3 allele (OR 7.7 vs. 6.8; P = 0.29; Table 2). That the apparent difference in risk for DRB1*03:01 haplotypes bearing DRB3*02:02 versus DRB3*01:01 is not evident in DRB1*03:01/DRB1*04 heterozygotes may reflect the very high risk associated with this genotype and attributed, in part, to the *trans*-complementing DQ-α (*05:01) or DQ-β (*03:02) heterodimer (4).

Although many samples were sequenced in this case/control study, the number of informative haplotypes (DRB1*03:01) and genotypes was limited; therefore, the statistical power of the association with DRB3:02:02 was very modest. Access to the T1DGC collection of HLA-genotyped families allowed selective targeting of informative genotypes to directly address the hypothesis of an effect of DRB3*02:02 in DRB1*03:01/*03:01 homozygotes and replicate the results of the case/control study.
Targeted replication study of informative samples from the T1DGC family collection. To further evaluate the putative role of *HLA-DRB3* polymorphism on DRB1*03:01 haplotype risk, a panel of DRB1*03:01/*03:01 homozygotes (n = 54 T1D family members, 105 nonshared chromosomes) and DRB1*03:01/DRB1*04-DQB1*03:02 (n = 227) heterozygous T1D members were selected from the T1DGC families. Of the DRB1*03:01/DRB1*04 subjects, 154 were *04:01 and 73 were *04:04. The distribution of DRB3 alleles in these T1D subjects is reported in Table 3. The frequency of the DRB3*02:02 allele in the DRB1*03:01/*03:01 T1D members was significantly greater than in those with DRB1*03:01/*04-DQB1*03:02 (42.9 vs. 27.6%, P < 0.005), consistent with the observations in the previous case/control study (Table 2). The distribution of DRB3 alleles on control DRB1*03:01 chromosomes was evaluated by examining the nontransmitted AFBAC (16) DRB1*03:01 chromosomes from heterozygous parents. The frequency of DRB3*02:02 in those chromosomes not transmitted to a T1D-affected child was 22.6%, significantly less than the 42.9% (P = 0.001) observed in the DRB1*03:01/DRB1*03:01 homozygous patients but not significantly different from the frequency in DRB1*03:01/DRB1*04:01 heterozygous T1D patients (27.3%).
Does DRB3*02:02 allele modify risk or is it simply a marker for a high-risk DRB1*03:01 haplotype? On the basis of these two studies, the DRB3*02:02 allele appears to mark a higher-risk DRB1*03:01 haplotype than the DRB1*03:01-DRB3*01:01 haplotype, and the effect of this high-risk haplotype can be discerned primarily in the DRB1*03:01/*03:01 homozygotes. T1D risk heterogeneity in DRB1*03:01 haplotypes has been previously reported with HLA-A*30:02, HLA-B*18:01, and DPB1*02:02 alleles

TABLE 1
DRB1-DRB3 and DRB1-DRB5 haplotype frequencies in T1D case and control subjects

DRB1	DRB3	DRB5	Control subjects (n = 996) n (%)	T1D (n = 984) n (%)	OR (95% CI)	P value
01:01			87 (8.7)	86 (8.7)	1.00 (0.73–1.37)	1.00
01:02			9 (0.9)	6 (0.6)	0.67 (0.24–1.90)	0.45
01:03			13 (1.3)	4 (0.4)	0.31 (0.10–0.95)	**3.1E-02**
03:01	**01:01**		**119 (11.9)**	**255 (25.9)**	**2.58 (2.03–3.27)**	**8.7E-13**
03:01	**02:02**		**24 (2.4)**	**82 (8.3)**	**3.68 (2.32–5.85)**	**1.2E-08**
04:01			122 (12.2)	270 (27.4)	2.71 (2.14–3.43)	**3.1E-14**
04:02			2 (0.2)	18 (1.8)	9.26 (2.14–40.0)	**3.1E-04**
04:04			56 (5.6)	82 (8.3)	1.53 (1.07–2.17)	**2.2E-02**
04:05			6 (0.6)	26 (2.6)	4.48 (1.84–10.9)	**3.6E-04**
04:07			11 (1.1)	2 (0.2)	0.18 (0.04–0.82)	**1.3E-02**
04:08			6 (0.6)	4 (0.4)	0.67 (0.19–2.39)	0.54
07:01			128 (12.9)	37 (3.8)	0.26 (0.18–0.39)	**2.4E-12**
08:01			28 (2.8)	20 (2.0)	0.72 (0.40–1.28)	0.27
09:01			7 (0.7)	8 (0.8)	1.16 (0.42–3.21)	0.78
10:01			5 (0.5)	1 (0.1)	0.20 (0.02–1.73)	0.11
11:01	02:02		35 (3.5)	11 (1.1)	0.31 (0.16–0.61)	**4.7E-04**
11:03	02:02		5 (0.5)	1 (0.1)	0.20 (0.02–1.73)	0.11
11:04	02:02		18 (1.8)	2 (0.2)	0.11 (0.03–0.48)	**3.8E-04**
12:01	02:02		19 (1.9)	6 (0.6)	0.32 (0.13–0.79)	**1.0E-02**
13:01	**01:01**		**32 (3.2)**	**5 (0.5)**	**0.15 (0.06–0.40)**	**1.1E-05**
13:01	**02:02**		**17 (1.7)**	**5 (0.5)**	**0.29 (0.11–0.80)**	**1.1E-02**
13:02	03:01		42 (4.2)	36 (3.7)	0.86 (0.55–1.36)	0.53
13:03	01:01		10 (1.0)	0 (0.0)	0.00	**1.7E-03**
14:01/54	02:02		28 (2.8)	1 (0.1)	0.04 (0.00–0.26)	**6.3E-07**
15:01		01:01	137 (13.8)	2 (0.2)	0.01 (0.00–0.05)	**5.3E-30**
16:01		02:02	2 (0.2)	4 (0.4)	2.03 (0.37–11.1)	0.41

Boldface type indicates statistical significance and DRB1 haplotypes (DRB1*03:01 and *13:01) with two different DRB3 alleles.

distinguishing the higher-risk from the lower-risk DRB1*03:01 haplotypes (17). To investigate whether the DRB3*02:02 allele might contribute to T1D risk or simply mark a high-risk DRB1*03:01 haplotype in which HLA class I (or other alleles) was a modifying risk, we examined the distribution of these and other T1D-associated alleles at other HLA loci on these two (DRB3*01:01 and DRB3*02:02) DRB1*03:01 haplotypes. The increase of DRB3*02:02 on DRB1*03:01 haplotypes in DRB1*03:01/*03:01 case subjects versus DRB1*03:01/DRB1*04 case subjects and AFBAC control subjects could reflect LD with alleles at other high-risk loci. The analysis of the case subjects selected from T1DGC families with informative HLA-DRB1 genotypes included parents; thus, haplotypes across the HLA region could be determined.

Eight locus haplotypes were phased and assigned based on familial inheritance. In the absence of parental genotyping data for the DRB3 locus, only a fraction of these could be assigned phase unambiguously with regard to the remaining eight loci. In total, 327 DRB1*03:01 haplotypes were used for this analysis of extended HLA haplotypes. The DRB1*03:01 haplotype counts in the three groups (AFBAC, DRB1*03:01/DRB1*04, DRB1*03:01/DRB1*03:01) for selected DPB1, HLA-B, -C, and -A alleles are reported in Table 4 and Supplementary Table 1. These alleles were selected based on previously published reports of T1D association. Table 4 and Supplementary Table 1 compare the distribution of DRB3*01:01 and *02:02 alleles on DRB1*03:01 haplotypes bearing various high-risk alleles at other HLA loci in DRB1*03:01/*03:01 versus DRB1*03:01/*04 patients.

We note that the alleles A*01:01, B*08:01, C*07:02, and DPB1*01:01 are in very strong LD with DRB3*01:01. These

TABLE 2
Association of DRB1*03:01/03:01 and DRB1*03:01/04:01 genotypes with T1D stratified by DRB3 genotype

DRB1 genotype	DRB3	Control subjects n (%)	T1D n (%)	OR (95% CI)	Z statistic	P value
03:01/03:01	02:02+	1 (0.2)	28 (5.7)	25.49 (3.43–189.2)	2.13	**0.033**
03:01/03:01	01:01/01:01	7 (1.4)	23 (4.67)	3.44 (1.46–8.09)		
03:01/04:01	02:02	3 (0.6)	22 (4.5)	7.72 (2.3–25.9)	1.06	0.290
03:01/04:01	01:01 or 01:02	14 (2.8)	81 (16.5)	6.81 (3.81–12.2)		

Boldface type indicates statistical significance. To assess the statistical significance of the difference between ORs for DRB3*02:02+ vs. DRB3*01:01/01:01 genotypes, the P value for a Z statistic was computed as $Z = |LnORi - LnORj|/\sqrt{(SE2i - SE2j)}$ where SEi denotes the standard error of the log OR LnORi, and i and j refer to DRB3*02:02+ vs. DRB3*01:01/010×. For P values estimated with the Yates continuity correction, $P = 0.064$ ($\chi^2 = 3.4228$, df = 1) for DR3/DR3/. $P = 0.052$ estimated using the Fisher exact test (two-sided). $P = 0.979$ for DR3/DR4 using the Yates continuity correction.

TABLE 3

Comparison of distribution of DRB3 alleles on DR3 haplotypes in control subjects (AFBAC), DRB1*03:01 homozygous patients, and DRB1*03:01/DRB1*04 heterozygous patients

DRB1 genotype	DRB3*01:01 allele n (%)	DRB3*02:02 allele n (%)	P value vs. DR3/DR4	P value vs. AFBAC
DR3/DR3	60 (57.1)	45 (42.9)	0.0048[a]	0.0013[b]
DR3/DR4[c]	165 (72.7)	62 (27.3)		0.347 ns
DR3 AFBAC	89 (77.4)	26 (22.6)	0.347 ns	

[a]$P < 0.0072$ with the Yates continuity correction. [b]$P < 0.0022$ with the Yates continuity correction. [c]The DRB1*04 designation in this table is the sum of the DRB1*04:01 and *04:04 alleles. ns, nonsignificant.

alleles on this extended haplotype mark a "lower-risk" DRB1*03:01 haplotype. In addition, the LD between alleles on the extended DRB1*03:01 haplotype, A*30:02-B*18:01-DRB1*03:01-DRB3*02:02-DPB1*02:02 is sufficiently strong that it does not permit evaluation of the independent contribution of DRB3*02:02. One haplotype, the DRB1*03:01-DPB1*03:01, however, is informative; DRB1*03:01 haplotypes carrying both DPB1*03:01 and DRB3*02:02 are significantly increased ($P = 0.0006$) in DRB1*03:01/*03:01 versus DRB1*03:01/*04 case subjects compared with those who carry DRB3*01:01 (Table 4). This effect cannot be explained by the presence of A*30:02 or B*18:01, which are also markers of high-risk DRB1*03:01 haplotypes on these 16 DPB1*03:01-DRB3*02:02 haplotypes. Only 4 of these 16 haplotypes carry A*30:02, and of the 9 DPB1*03:01-DRB3*02:02 haplotypes that also carry B*18:01, 5 are found in DR3/DR4 and 4 in DR3/DR3 case subjects. These observations, taken together, suggest that the DRB3*02:02 allele may contribute to T1D risk rather than simply marking a high-risk DRB1*03:01 haplotype.

DRB1*13:01 is the only other DRB1 haplotype that can carry DRB3*01:01 or DRB3*02:02 (Table 1). We note that the OR of the DRB1*13:01-DRB3*02:02 haplotype is nominally higher than the DRB1*13:01-DRB3*01:01 haplotype (0.29 vs. 0.15; Table 1). However, the CIs overlap, indicating that much larger sample sizes will be necessary to evaluate the risk of these two haplotypes.

DISCUSSION

Long-read next generation clonal sequencing allows genotyping the secondary HLA-DRB loci (*HLA-DRB3*, *HLA-DRB4*, and *HLA-DRB5*) as well as *HLA-DRB1* by using DRB generic 454 fusion primers for exon 2. Using the Roche 454 GS FLX and GS Junior Systems, we have used this capability to examine the role of the secondary HLA-DRB loci in T1D susceptibility. Most *HLA-DRB1* alleles are linked exclusively to a specific allele at the secondary DRB locus (Table 1); however, DRB1*03:01 and DRB1*13:01 can carry DRB3*01:01 or *02:02. DRB1*03:01 haplotypes carrying DRB3*02:02 appear to confer higher risk than

those carrying DRB3*01:01. The difference in T1D risk is observed in DRB1*03:01/*03:01 homozygotes but not in DRB1*03:01/DRB1*04-DQB1*03:02 heterozygotes. The very high risk for T1D associated with this heterozygous genotype has been attributed to the *trans*-complementing DQ heterodimer formed by the DQ-α chain encoded by the DQA1*05:01 allele on the DRB1*03:01 haplotype and the DQ-β chain encoded by the DQB1*03:02 on the DRB1*04-DQB1*03:02 haplotype (4). One possible explanation of the different effect of *HLA-DRB3* in this heterozygote and in the DRB1*03:01/DRB1*03:01 homozygote is that the putative risk conferred by the *trans*-complementing DQ heterodimers in the heterozygote is sufficiently large so that the risk difference between the "higher-risk" and "lower-risk" DRB1*03:01 haplotypes has a minimal effect on the overall T1D risk for the heterozygote. A recent study on the effect of other MHC markers (*HLA-B*, *HLA-A*, *HLA-DPB1*, and *TNF-α*) on DRB1*03:01 haplotype risk reported a similar observation—these markers were associated with differential risk in DRB1*03:01/DRB1*03:01 homozygotes but not in the DRB1*03:01/DRB1*04-DQB1*03:02 heterozygotes (17).

Does DRB3*02:02 play a role in T1D risk or does it only *mark* the higher-risk DRB1*03:01 haplotype? To address this question, eight locus DRB1*03:01 haplotypes were phased and assigned based on familial inheritance, and the distribution of the DRB3*01:01 and *02:02 alleles on extended haplotypes bearing high-risk alleles at other HLA loci was compared (Table 4 and Supplementary Table 1). The linkage disequilibrium between A*30:02, B*18:01, DPB1*02:02, and DRB3*02:02 was so great (virtually 100%) that the effect of these alleles on this high-risk DRB1*03:01 haplotype could not be distinguished. Some DRB1*03:01-DRB3*02:02 haplotypes, however, *do not* carry these alleles, associated with the high-risk DRB1*03:01 haplotype, but still demonstrate a significant increase in DRB1*03:01/*03:01 case subjects versus DRB1*03:01/*04. In particular, DRB1*03:01 haplotypes bearing DPB1*03:01 can carry DRB3*01:01 or DRB3*02:02; those that carry DRB3*02:02 are dramatically increased among the DRB1*03:01/*03:01 case subjects ($P = 0.0006$). These observations are

TABLE 4

Distribution of DRB3*02:02 and *01:01 alleles on the extended DRB1*03:01-DPB1*03:01 haplotype

		AFBAC n (%)	DR3/4 n (%)	DR3/3 n (%)	T1D vs. control subjects P <	DR3/3 vs. DR3/4 T1D P <
DPB1*03:01	DRB3*02:02	1 (25.0)	7 (18.4)	8 (72.7)	0.81	0.0006
DPB1*03:01	DRB3*01:01	3 (75.0)	31 (81.6)	3 (27.3)		

With the Yates correction, $P < 0.74$ and $P < 0.0021$, respectively.

consistent with the hypothesis that DRB3*02:02 is not simply a marker of high-risk haplotypes but, in fact, increases the risk of DRB1*03:01 haplotypes, particularly in DRB1*03:01/*03:01 homozygotes.

The amino acid sequence differences between DRB3*01:01 and DRB3*02:02 are substantial (10 differences). Consequently, DRB3*02:02 is likely to have a different repertoire of peptide binding and presentation and may thus confer greater T1D risk than DRB3*01:01 in certain genotype contexts (i.e., DRB1*03:01/*03:01 homozygotes) through altered peptide binding and T-cell repertoire. The crystallographic structure of the DRB3*01:01 (DR52a) and the DRB3*03:01 (DR52c) molecules have been reported recently (18,19). The P9 pocket of the DRB3*02:02 (DR52b) differs substantially from that of DRB3*01:01 (DR52a); in particular, it contains Tyr-37, Ala-38, Asp-57, and Tyr-60, making the pocket more accommodating to smaller, polar, or charged peptide residues. A structural model with bound peptide of the DRB3*02:02 molecule with the differences highlighted is shown in Supplementary Fig. 1.

The results presented here support the conclusion that the T1D risk of a given HLA haplotype is determined by specific combinations of alleles at a variety of HLA loci with genotype-dependent effects and support a role for the DRB3 locus in T1D susceptibility conferred by the DRB1*03:01 haplotype.

REFERENCES

1. Rich SS, Concannon P, Erlich H, et al. The Type 1 Diabetes Genetics Consortium. Ann N Y Acad Sci 2006;1079:1–8
2. Concannon P, Erlich HA, Julier C, et al.; Type 1 Diabetes Genetics Consortium. Type 1 diabetes: evidence for susceptibility loci from four genome-wide linkage scans in 1,435 multiplex families. Diabetes 2005;54:2995–3001
3. Barrett JC, Clayton DG, Concannon P, et al.; Type 1 Diabetes Genetics Consortium. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 2009;41:703–707
4. Erlich H, Valdes AM, Noble J, et al.; Type 1 Diabetes Genetics Consortium. HLA DR-DQ haplotypes and genotypes and type 1 diabetes risk: analysis of the type 1 diabetes genetics consortium families. Diabetes 2008;57:1084–1092
5. Noble JA, Valdes AM, Varney MD, et al.; Type 1 Diabetes Genetics Consortium. HLA class I and genetic susceptibility to type 1 diabetes: results from the Type 1 Diabetes Genetics Consortium. Diabetes 2010;59:2972–2979
6. Varney MD, Valdes AM, Carlson JA, et al.; Type 1 Diabetes Genetics Consortium. HLA DPA1, DPB1 alleles and haplotypes contribute to the risk associated with type 1 diabetes: analysis of the type 1 diabetes genetics consortium families. Diabetes 2010;59:2055–2062
7. Howson JM, Walker NM, Clayton D, Todd JA; Type 1 Diabetes Genetics Consortium. Confirmation of HLA class II independent type 1 diabetes associations in the major histocompatibility complex including HLA-B and HLA-A. Diabetes Obes Metab 2009;11(Suppl. 1):31–45
8. Valdes AM, Erlich HA, Noble JA. Human leukocyte antigen class I B and C loci contribute to type 1 diabetes (T1D) susceptibility and age at T1D onset. Hum Immunol 2005;66:301–313
9. Nejentsev S, Howson JM, Walker NM, et al.; Wellcome Trust Case Control Consortium. Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. Nature 2007;450:887–892
10. Noble JA, Valdes AM, Bugawan TL, Apple RJ, Thomson G, Erlich HA. The HLA class I A locus affects susceptibility to type 1 diabetes. Hum Immunol 2002;63:657–664
11. Noble JA, Valdes AM, Thomson G, Erlich HA. The HLA class II locus DPB1 can influence susceptibility to type 1 diabetes. Diabetes 2000;49:121–125
12. Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature 2005;437:376–380
13. Bentley G, Higuchi R, Hoglund B, et al. High-resolution, high-throughput HLA genotyping by next-generation sequencing. Tissue Antigens 2009;74:393–403
14. Holcomb CL, Höglund B, Anderson MW, et al. A multi-site study using high-resolution HLA genotyping by next generation sequencing. Tissue Antigens 2011;77:206–217
15. Todd JA, Walker NM, Cooper JD, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet 2007;39:857–864
16. Thomson G. Mapping disease genes: family-based association studies. Am J Hum Genet 1995;57:487–498
17. Noble JA, Martin A, Valdes AM, et al. Type 1 diabetes risk for human leukocyte antigen (HLA)-DR3 haplotypes depends on genotypic context: association of DPB1 and HLA class I loci among DR3- and DR4-matched Italian patients and controls. Hum Immunol 2008;69:291–300
18. Parry CS, Gorski J, Stern LJ. Crystallographic structure of the human leukocyte antigen DRA, DRB3*0101: models of a directional alloimmune response and autoimmunity. J Mol Biol 2007;371:435–446
19. Dai S, Crawford F, Marrack P, Kappler JW. The structure of HLA-DR52c: comparison to other HLA-DRB3 alleles. Proc Natl Acad Sci USA 2008;105:11893–11897