Published in final edited form as:

Diabetes Obes Metab. 2009 February ; 11(Suppl 1): 25–30. doi:10.1111/j.1463-1326.2008.01000.x.

The frequent and conserved DR3-B8-A1 extended haplotype confers less diabetes risk than other DR3 haplotypes

E. E. Baschal^{1,2}, T. A. Aly^{1,2}, J. M. Jasinski^{1,2}, A. K. Steck¹, K. N. Johnson¹, J. A. Noble³, H. A. Erlich⁴, G. S. Eisenbarth^{1,2}, and the Type 1 Diabetes Genetics Consortium

¹Barbara Davis Center for Childhood Diabetes, Department of Pediatrics, University of Colorado Denver, Aurora, CO, USA

²Human Medical Genetics Program, University of Colorado Denver, Aurora, CO, USA

³Children's Hospital Oakland Research Institute, Oakland, CA, USA

⁴Roche Molecular Systems, Alameda, CA, USA

Abstract

Aim—The goal of this study was to develop and implement methodology that would aid in the analysis of extended high-density single nucleotide polymorphism (SNP) major histocompatibility complex (MHC) haplotypes combined with human leucocyte antigen (HLA) alleles in relation to type 1 diabetes risk.

Methods—High-density SNP genotype data (2918 SNPs) across the MHC from the Type 1 Diabetes Genetics Consortium (1240 families), in addition to HLA data, were processed into haplotypes using PEDCHECK and MERLIN, and extended DR3 haplotypes were analysed.

Results—With this large dense set of SNPs, the conservation of DR3-B8-A1 (8.1) haplotypes spanned the MHC ($\geq 99\%$ SNP identity). Forty-seven individuals homozygous for the 8.1 haplotype also shared the same homozygous genotype at four 'sentinel' SNPs (rs2157678 'T', rs3130380 'A', rs3094628 'C' and rs3130352 'T'). Conservation extended from *HLA-DQB1* to the telomeric end of the SNP panels (3.4 Mb total). In addition, we found that the 8.1 haplotype is associated with lower risk than other DR3 haplotypes by both haplotypic and genotypic analyses [haplotype: $p = 0.009$, odds ratio (OR) = 0.65; genotype: $p = 6.3 \times 10^{-5}$, OR = 0.27]. The 8.1 haplotype (from genotypic analyses) is associated with lower risk than the high-risk DR3-B18-A30 haplotype ($p = 0.01$, OR = 0.23), but the DR3-B18-A30 haplotype did not differ from other non-8.1 DR3 haplotypes relative to diabetes association.

Conclusion—The 8.1 haplotype demonstrates extreme conservation (>3.4 Mb) and is associated with significantly lower risk for type 1 diabetes than other DR3 haplotypes.

Keywords

8.1 haplotype; extended haplotypes; major histocompatibility complex; T1DGC; type 1 diabetes

© 2009 The Authors Journal Compilation © 2009 Blackwell Publishing Ltd

Correspondence: George S. Eisenbarth, MD, PhD, Barbara Davis Center for Childhood Diabetes, University of Colorado Denver, Mail Stop B140, PO Box 6511, Aurora, CO 80045-6511, USA. george.eisenbarth@UCHSC.edu.

Conflict of interest:

The authors do not declare any conflict of interest relevant to this manuscript.

Please note: Wiley-Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Introduction

It is now possible to rapidly and relatively inexpensively analyze thousands of single nucleotide polymorphisms (SNPs) across any section of the genome including the Major Histocompatibility Complex (MHC). In particular, custom SNP panels have been developed that aid in the analysis of the MHC. Most MHC analyses to date have evaluated disease associations of haplotype blocks ranging in size from 5000 to 200 000 base pairs [1]. It has long been known that within the MHC, there are extended haplotypes that include the classical MHC region telomeric of human leucocyte antigen (*HLA*)-*DP*, which with current SNP typing can be identified to include regions of almost total identity ($\geq 99\%$), often greater than 2.9 million base pairs [1,2]. A seminal study 25 years ago analysed HLA alleles (*HLA-B* and *HLA-DR*) and polymorphisms in complement genes (complotypes, a single genetic unit of the complement genes *CFB*, *C2*, *C4A* and *C4B*) and identified multiple extended haplotypes across this region [3]. This report has been confirmed in several subsequent studies, with haplotypes always defined by HLA alleles and/or complotypes [1,4–6]. In addition, recent studies have confirmed that for the most common extended haplotypes (e.g., HLA-DR3-B8-A1; DR3-B18-A30), nearly complete conservation between unrelated individuals for up to 9 million base pairs can be found when SNPs are analysed in conjunction with HLA alleles [2,7–9]. The most common extended haplotype in patients with diabetes, the *HLA-DRB1**03 (DR3), *HLA-B**08 (B8) and *HLA-A**01 (A1) (8.1), is present in 9% of Caucasian MHC control haplotypes and 18% of MHC case haplotypes from families with type 1 diabetes [1]. The 8.1 haplotype has been previously reported to be associated with lower risk for type 1 diabetes than the high-risk DR3-B18-A30 (Basque haplotype) [7,10] and slightly lower risk than other DR3 haplotypes and has been associated with a series of autoimmune diseases [11].

The Type 1 Diabetes Genetics Consortium (T1DGC) performed high-density SNP typing of the classical MHC region using two sets of standard Illumina™ SNP panels (Exon-centric and Mapping, 3072 SNPs total, Illumina Corporation, San Diego, CA, USA) and tested 1240 families with at least two siblings affected by type 1 diabetes. Using these data, we have selected and utilized a set of analysis packages to define MHC haplotypes in these families. The accuracy of the typing and analyses in forming haplotypes is reflected with not only standard Mendelian inheritance verification but also in the initial analysis of the most common extended HLA haplotype (8.1). Of note, this haplotypic analysis indicates that the 8.1 haplotype of DR3-bearing chromosomes is associated with decreased risk for type 1 diabetes relative to non-8.1 DR3-bearing chromosomes.

Materials and Methods

This study included 1240 families (6297 individuals, mostly affected sib pairs and their parents) from the British Diabetic Association, Danish, Joslin, Human Biological Data Interchange and UK populations. SNPs were typed across the MHC using the dense standard Illumina MHC Mapping (#GT-17–181) and Exon-centric (#GT-17–191) panels [2957 distinct SNPs (1536 SNPs in each panel with 115 overlapping SNPs) with 2837 of 2957 SNPs successfully typed, yielding a 96% SNP success rate]. In addition, complete HLA typing (*HLA-DPB1*, *HLA-DPA1*, *HLA-DQB1*, *HLA-DQA1*, *HLA-DRB1*, *HLA-B*, *HLA-C* and *HLA-A*, performed using a traditional PCR-based sequence-specific oligonucleotide probe system where oligonucleotide probes were immobilized on nylon membranes [12]) was available.

Haplotypes were generated from SNP genotype data using the following method. First, to establish that the genotype data demonstrated a Mendelian inheritance pattern within each family, the PEDCHECK program [13] was used (<http://watson.hgen.pitt.edu>) on data from both Illumina panels and HLA separately. Mendelian inheritance patterns were present for all families. Next, data from the Mapping SNP panel, the Exon-centric SNP panel and HLA were

combined. One hundred and fifteen SNPs were present in both the Mapping panel and the Exon-centric panel. For these duplicate SNPs, the result (from either the Mapping panel or the Exon-centric panel) with the lower success rate was excluded from the analysis. MERLIN software (www.sph.umich.edu/csg/abecasis/Merlin) [14] was used to phase the SNP genotype data from families into haplotypes. AFBAC (affected family-based control) methodology was used to assign case or control status to chromosomes [15] using a VBA (Visual Basic for Applications) macro (in Microsoft Excel™, Microsoft Corporation, Redmond, WA, USA) created “in house” (Appendix S1). Microsoft Excel 2007 was used to process and analyze chromosome data. Individuals with ‘unknown’ type 1 diabetes status in the database were excluded from these analyses. Conserved extended 8.1 haplotypes were identified by the presence of HLA-DR3, -B8 and -A1 alleles and phased data for at least two of the four sentinel 8.1 SNPs (SNPs selected to be highly associated with the 8.1 extended haplotype: rs2157678 ‘T’, rs3130380 ‘A’, rs3094628 ‘C’ and rs3130352 ‘T’) and none of the four SNPs discordant with the defined 8.1 extended haplotype for both chromosome studies and analyses of 8.1 homozygous individuals. Chi-squared and Fisher’s exact tests were performed using an alpha level for significance of 0.05.

Results

Haplotypes were generated as described in the *Materials and Methods* using MERLIN to phase alleles. MERLIN does not infer phase when phase cannot be determined (if both the parents and the child are heterozygous for a SNP for example), in contrast to the commonly used program PHASE (www.stat.washington.edu/stephens/software.html), which uses statistical analysis of allele frequencies to indicate likely phase. Therefore, when phase is indicated in MERLIN, it is unambiguous. In this data set, using MERLIN to create haplotypes, only 4% of analysed SNPs are unphased.

We examined only founder (parental) chromosomes (each of four founder chromosomes in a family are represented only once in the data set) for all the following chromosome analyses. Within founder chromosomes of all DR types, DR3 was present at a frequency of 32% in case chromosomes (898/2786) and 12% in control chromosomes (176/1528) [$p = 1.3 \times 10^{-55}$, odds ratio (OR) = 3.7 (95% confidence interval 3.1–4.4)], confirming the higher risk associated with DR3.

For the chromosome studies, we defined the conserved 8.1 haplotype as having HLA alleles DR3, B8 and A1, and in addition having specific alleles at four sentinel 8.1 SNPs (see *Materials and Methods*). Within founder chromosomes of all DR types, the conserved 8.1 haplotype was present in 13% of case chromosomes (357/2786) and 6% of control chromosomes (89/1528) [$p = 7.1 \times 10^{-14}$, OR = 2.4 (1.9–3.0)]. Among DR3 chromosomes (table 1), the conserved 8.1 haplotype is present in 40% of case chromosomes (357/898) and 50% of control chromosomes (89/176) [$p = 9.4 \times 10^{-3}$, OR = 0.65 (0.47–0.89)]. Past studies have identified DR3-B18-A30 (also known as the ‘Basque’ haplotype because of its high frequency in Basque populations) as a particularly high-risk DR3 haplotype [7]. In this analysis, the DR3-B18-A30 haplotype is associated with higher risk than the conserved 8.1 haplotype [DR3-B18-A30: 59 cases, 5 controls; 8.1: 357 cases, 89 controls; $p = 0.02$, OR = 2.9 (1.2–7.6)].

To illustrate the regions where conservation of the 8.1 haplotype decays, figure 1A–C depict haplotypes of the 47 DR3 8.1 homozygous individuals (to allow the figure to have only three panels, the intervening ‘conserved’ region between panels A and B is not shown). Each column is a chromosome (one individual per family is represented), and each row is a SNP. The extent of almost complete conservation is apparent. The chromosomes are ordered by length of conservation from *HLA-DQB1* telomeric. At the centromeric end, near *HLA-DPBI*, conservation is ‘acutely’ lost for approximately half of the chromosomes. At the telomeric end,

most chromosomes remained conserved to the end of the SNP panel. Two of the 94 chromosomes had only 'limited' long regions of 8.1 identity despite selection for HLA-DR3, -B8 and -A1 alleles and 8.1 matching alleles at four 8.1 sentinel SNPs, which are located just telomeric (30.24–30.44 Mb) of the regions that lack identity for these two haplotypes (on the far right of figure 1B).

Next, to look at the effect of the 8.1 haplotype in individuals, we examined 278 DR3/3 homozygotes (200 cases and 78 controls) that included 186 individuals with an 8.1 haplotype (120 cases and 66 controls). We compared genotypes for individuals with an 8.1 haplotype vs. non-8.1 DR3/3 individuals (table 2). As seen from table 2, when DR3/3 individuals are compared across the three genotypes (homozygous 8.1, heterozygous 8.1 and homozygous non-8.1), there is a significant difference between the three groups (Chi-squared $p = 4 \times 10^{-4}$). When individuals with at least one copy of the 8.1 haplotype ($N = 120$ cases and 66 controls) were compared with non-8.1 DR3/3 individuals ($N = 80$ cases and 12 controls), the 8.1 haplotype is significantly less associated with type 1 diabetes than non-8.1 DR3 haplotypes [$p = 6.3 \times 10^{-5}$, OR = 0.27 (0.14–54)] (heterozygous 8.1 DR3/3 did not differ from homozygous 8.1 in diabetes association, $p = 0.9$).

To confirm that the protective effect of the 8.1 haplotype was not solely because of the inclusion of the high-risk DR3-B18-A30 haplotype in the non-8.1 DR3/3 group, we excluded 43 individuals with one or more DR3-B18-A30 haplotypes (table 2). In this case, individuals with at least one copy of the 8.1 haplotype are at lower risk for type 1 diabetes than DR3/3 individuals with no copies of the 8.1 haplotype [8.1 haplotype: 113/168 (67%) of DR3/3 cases and 58/67 (86%) of DR3/3 controls; $p = 3.2 \times 10^{-3}$, OR = 0.32 (0.15–0.69)]. This confirms that the 8.1 haplotype is associated with lower risk than other DR3 haplotypes and that this effect is not because of the inclusion of the high-risk DR3-B18-A30 haplotype in the non-8.1 DR3/3 group.

Among homozygous DR3/3 individuals when individuals with at least one copy of the 8.1 haplotype (113 cases and 58 controls) are compared with individuals with at least one copy of the high-risk DR3-B18-A30 haplotype (25 cases and 3 controls), the 8.1 haplotype is protective [$p = 0.01$, OR = 0.23 (0.07–0.81)]. Additionally, we wanted to examine the effect of the DR3-B18-A30 haplotype (table 2) and found that when individuals with an 8.1 haplotype are excluded, individuals with at least one copy of the DR3-B18-A30 haplotype were not at a significantly higher risk compared with non-DR3-B18-A30 DR3/3 individuals [$p = 0.75$, OR = 1.36 (0.34–5.47)]. This may indicate that the previously observed increased risk associated with the DR3-B18-A30 haplotype is because of the protective effect of the 8.1 DR3 haplotype.

Discussion

Having large data sets with dense SNP typing of complete families allows definitive generation of haplotypes across the MHC. In the current communication, we have analyzed data from 1240 affected sib-pair families, giving us the ability to examine the conserved 8.1 haplotype on chromosomes composed of both HLA and SNP data. Standard analysis programs (PEDCHECK and MERLIN) with current PC computers allow rapid (less than 24 hours) generation of complete haplotypes despite the size of the data sets provided by the MHC fine mapping Illumina panels. Figure 1 depicts the extensive SNP conservation among individuals homozygous for the DR-B8-A1 haplotype, which is the most common extended MHC haplotype within Caucasian populations. Overall (compared with all MHC haplotypes, not only those with a DR3), this conserved haplotype is associated with increased type 1 diabetes risk, presumably associated with its DR3-DQB1*0201 alleles. The current study indicates that there is a dramatic decrease in association with type 1 diabetes for DR3/3 individuals with at least one copy of the 8.1 haplotype compared with those individuals homozygous for DR3 but lacking the 8.1 haplotype. For instance, 15% of control individuals are homozygous for non-8.1

DR3 haplotypes vs. 40% of the family members with type 1 diabetes. The protective effect of the 8.1 haplotype vs. other DR3 haplotypes is probably underestimated in the current data set in those families that were selected to have two siblings with type 1 diabetes and control DR3 homozygous individuals include parents where an 8.1 haplotype is present in their offspring with diabetes.

The Basque extended DR3 haplotype (DR3-B18-A30) has been reported to be of higher risk compared with other DR3 haplotypes [7], and it has extended SNP identity (although very different SNP patterns compared with the 8.1 haplotype). In the current set of T1DGC DR3/3 individuals, the DR3-B18-A30 haplotype, when present in one or two copies ($N = 28$), did not confer significantly higher risk for type 1 diabetes when compared with non-8.1 DR3/3 individuals [$p = 0.75$, $OR = 1.4$ (0.3–5.5)], although with larger numbers, perhaps this haplotype would show increased association. The dramatic difference between 8.1 and non-8.1 DR3–bearing haplotypes indicates that non-DR/DQ loci, likely in the large region of 8.1 conservation, modify the risk conferred by DR3-DQB1*0201 alleles. Our data indicate that a single copy of the 8.1 haplotype in DR3/3 homozygous individuals confers similar protection as two copies (a dominant protective effect). We believe that the analysis of extended haplotypes will be important in defining the risk of diabetes provided by the MHC and will contribute to the identification of MHC region non-class II modifiers of type 1 diabetes risk. However, localizing causative genes may be more complicated than anticipated because of the long-range conservation inherent to the MHC region.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research utilizes resources provided by the T1DGC, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Allergy and Infectious Diseases, National Human Genome Research Institute, National Institute of Child Health and Human Development and Juvenile Diabetes Research Foundation International and supported by U01 DK062418. We thank Elise Eller for bioinformatics assistance. This work was supported by the National Institutes of Health (DK32083 and DK057538), Diabetes Autoimmunity Study in the Young (DK32493), Autoimmunity Prevention Center (AI050864), Diabetes Endocrine Research Center (P30 DK57516), Clinical Research Centers (MO1 RR00069 and MO1 RR00051), the Immune Tolerance Network (AI15416), the American Diabetes Association, the Juvenile Diabetes Research Foundation, the Children's Diabetes Foundation and the Brehm Coalition.

References

1. Alper CA, Larsen CE, Dubey DP, Awdeh ZL, Fici DA, Yunis EJ. The haplotype structure of the human major histocompatibility complex. *Hum Immunol* 2006;67:73–84. [PubMed: 16698428]
2. Aly TA, Eller E, Ide A, et al. Multi-SNP analysis of MHC region: remarkable conservation of HLA-A1-B8-DR3 haplotype. *Diabetes* 2006;55:1265–1269. [PubMed: 16644681]
3. Awdeh ZL, Raum D, Yunis EJ, Alper CA. Extended HLA/complement allele haplotypes: evidence for T/t-like complex in man. *Proc Natl Acad Sci U S A* 1983;80:259–263. [PubMed: 6401863]
4. Yunis EJ. 1987 Philip Levine award lecture. MHC haplotypes in biology and medicine. *Am J Clin Pathol* 1988;89:268–280. [PubMed: 2963535]
5. Degli-Esposti MA, Leaver AL, Christiansen FT, Witt CS, Abraham LJ, Dawkins RL. Ancestral haplotypes: conserved population MHC haplotypes. *Hum Immunol* 1992;34:242–252. [PubMed: 1464552]
6. Yunis EJ, Larsen CE, Fernandez-Vina M, et al. Inheritable variable sizes of DNA stretches in the human MHC: conserved extended haplotypes and their fragments or blocks. *Tissue Antigens* 2003;62:1–20. [PubMed: 12859592]

7. Bilbao JR, Calvo B, Aransay AM, et al. Conserved extended haplotypes discriminate HLA-DR3-homozygous Basque patients with type 1 diabetes mellitus and celiac disease. *Genes Immun* 2006;7:550–554. [PubMed: 16929349]
8. Romero V, Larsen CE, Duke-Cohan JS, et al. Genetic fixity in the human major histocompatibility complex and block size diversity in the class I region including HLA-E. *BMC Genet* 2007;8:14. [PubMed: 17430593]
9. Aly TA, Baschal EE, Jahromi MM, et al. Analysis of single nucleotide polymorphisms identifies major type 1A diabetes locus telomeric of the major histocompatibility complex. *Diabetes* 2008;57:770–776. [PubMed: 18065518]
10. Kelly H, McCann VJ, Kay PH, Dawkins RL. Susceptibility to IDDM is marked by MHC supratypes rather than individual alleles. *Immunogenetics* 1985;22:643–651. [PubMed: 3865895]
11. Robinson WP, Barbosa J, Rich SS, Thomson G. Homozygous parent affected sib pair method for detecting disease predisposing variants: application to insulin-dependent diabetes mellitus. *Genet Epidemiol* 1993;10:273–288. [PubMed: 8224807]
12. Erlich H, Valdes AM, Noble J, et al. 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. [PubMed: 18252895]
13. O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998;63:259–266. [PubMed: 9634505]
14. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002;30:97–101. [PubMed: 11731797]
15. Thomson G. Mapping disease genes: family-based association studies. *Am J Hum Genet* 1995;57:487–498. [PubMed: 7668275]

Fig. 1.

Data from chromosomes from 47 individuals homozygous for the 8.1 haplotype (one individual per family is represented) were placed in columns in Microsoft Excel ($N = 94$ chromosomes), with each individual separated by a black line, with a row for each single nucleotide polymorphism (SNP) displayed ($N = 2827$; SNPs were removed if they failed in all the 8.1 homozygous individuals). To derive an '8.1 consensus sequence', the 8.1 haplotypes were analysed to identify the most frequent allele for each SNP. The allele for each SNP of each chromosome was compared with the consensus 8.1 SNP alleles. Yellow highlights SNP alleles identical to the consensus 8.1, blue highlights variant SNP alleles different from the consensus 8.1 and white highlights untyped or untyped SNPs. For space reasons, the figure has been divided into the following: (A) Only 29.30–30.20 Mb is shown ($N = 561$ SNPs). (B) Only 31.54–32.74 Mb is shown ($N = 698$ SNPs). (C) Only 32.74–34.24 Mb is shown ($N = 600$ SNPs).

Table 1

All DR3 chromosomes (N = 1074) subdivided by the presence of common DR3 extended haplotypes

	8.1 (DR3-B*8-A*1 + 4 SNPs), N (%)	DR3- B*18-A*30, N (%)	Other DR3, N (%)	Total DR3
Case	357 (40)	59 (6)	482 (54)	898
Control	89 (50)	5 (3)	82 (47)	176

SNP, single nucleotide polymorphism.

Table 2

DR3/3 individuals subdivided by presence or absence of extended conserved haplotypes

8.1 individuals compared with DR3/3 (N = 278), N (%)

	8.1/8.1[*]	3/8.1^{***}	3/3	8.1/8.1 and 3/8.1^{***}
Case	30 (15)	90 (45)	80 (40)	120 (60)
Control	15 (19)	51 (66)	12 (15)	66 (85)

8.1 individuals (exclude DR3-B18-A30) compared with DR3/3 (N = 235), N (%)

	8.1/8.1[*]	3/8.1^{**}	3/3	8.1/8.1 and 3/8.1^{**}
Case	30 (18)	83 (49)	55 (33)	113 (67)
Control	15 (23)	43 (64)	9 (13)	58 (87)

DR3-B18-A30 individuals (exclude 8.1) compared with DR3/3 (N = 92), N (%)

	B/B	3/B	3/3	B/B and 3/B
Case	2 (2)	23 (29)	55 (69)	25 (31)
Control	0 (0)	3 (25)	9 (75)	3 (25)

8.1, DR3-B8-A1; B, 'Basque' DR3-B18-A30.

*
p < 0.05 when compared with 3/3.**
p < 0.01 when compared with 3/3.***
p < 0.001 when compared with 3/3.