NIH Public Access
Author Manuscript
Diabetes Obes Metab. Author manuscript; available in PMC 2009 October 1.

Published in final edited form as:
Diabetes Obes Metab. 2009 February ; 11(Suppl 1): 46–52. doi:10.1111/j.1463-1326.2008.01002.x.

Several loci in the HLA class III region are associated with T1D risk after adjusting for DRB1-DQB1

A. M. Valdes[1], **G. Thomson**[2], and **Type 1 Diabetes Genetics Consortium**

[1]Twin Research Unit, King's College London, St Thomas' Hospital, London, UK

[2]Department of Integrative Biology, University of California, Berkeley, CA, USA

Abstract

Aim—Several studies have indicated that genes in the human leucocyte antigen (HLA) region additional to the HLA class II DRB1-DQB1 contribute to type 1 diabetes (T1D) susceptibility. The aim of this study was to assess if markers in the class III Major Histocompatibility Complex (MHC) region are associated with T1D after accounting for linkage disequilibrium (LD) with DRB1-DQB1.

Methods—We investigated 356 single nucleotide polymorphisms (SNPs) in the class III region covering 1.1 megabases in two subsets of data: 289 Human Biological Data Interchange (HBDI) Caucasian families and 597 additional Caucasian families collected by the Type 1 Diabetes Genetics Consortium (T1DGC). Analysis conditioning on DRB1-DQB1 was performed using the overall conditional genotype method.

Results—Thirteen SNPs replicated in both subsets of the data and showed evidence of an additional effect on disease risk. Although some of the SNPs are in tight LD with each other, at least six of the associations were not because of LD with other class III markers. The strongest association within class III markers was with rs2395106 that maps 5′ to the *NOTCH4* gene, which has also been implicated in susceptibility to rheumatoid arthritis. The second association was with rs707915 mapping to the *MSH5* gene, in a block of six markers significantly associated with T1D after adjusting for LD with DR-DQ. In total, six-independent associations within class III were observed although results were not adjusted for LD with class I.

Conclusions—Our data confirm that the class III region is involved in T1D susceptibility and suggest that more than one gene in the region is involved.

Keywords

association; HLA class III; MHC; risk; SNP; type 1 diabetes

Introduction

The direct involvement of the human leucocyte antigen (HLA) classical class II DR-DQ (DRB1 and DQB1) genes in type 1 diabetes (T1D) is well established. A hierarchy of risk effects is seen from very predisposing, through predisposing, neutral, protective and very protective at the allele, haplotype, and genotype levels (see 1 and references therein).

Correspondence: Ana M. Valdes, Twin Research Unit, King's College London, St Thomas' Hospital, Lambeth Palace Road, London SE1 7EH, UK. ana.valdes@kcl.ac.uk.

Conflict of Interest: The authors declare that they have no conflicts of interest in publishing this article.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

A number of analytic strategies have been developed to detect genes or markers additional to a primary predisposing gene in a genetic region; they all rely on stratification analyses to take account of the effects of linkage disequilibrium (LD) with the primary predisposing gene(s), that is, HLA DR-DQ in the case of T1D (for review see 2). Some use only a restricted set of the data. These include matched genotype strategies [3], and for family-based data, the homozygous parent transmission disequilibrium test (TDT) [4], and the homozygous parent linkage test [5].

The conditional haplotype method [6,7] and the conditional extended TDT [8] use all the data, but haplotypes need to be estimated. Basically, under the null hypothesis, the relative frequencies of the marker alleles on haplotypes stratified by disease locus alleles should be the same in cases and controls. While fit to these expectations does not exclude the possibility that other genes in the HLA complex are involved in disease, lack of fit unequivocally shows that all disease predisposing genes in the region have not been identified.

Similarly, the conditional genotype method examines the relative frequencies of genotypes at marker loci on specific HLA DR-DQ genotypes [9]. An obvious advantage of a conditional genotype method is that haplotype frequencies do not need to be estimated.

Thomson [10] extended the conditional haplotype approach with development of a method to test for additional genetic effects over all haplotypes; this is now referred to as the overall conditional haplotype method. In an analogous way, Thomson and Valdes [9] extended the conditional genotype method to test over all genotypes: the overall conditional genotype method (OCGM). Other methods to detect additional disease genes in a region are conditional logistic regression and methods combining association and linkage data [11].

Controlling for the influence of class II DR-DQ haplotype and genotype effects, a role in T1D has been shown of additional HLA class II (DPB1) [11,12] and class I genes including age-of-onset effects and rapid disease progression (13–15 and references therein). Also, various analyses have shown the presence of additional disease predisposing markers on specific high-risk DR-DQ haplotypes and genotypes, that is, DR3 and DR4 [3,16–19]. Possible heterogeneity of the DR15 haplotype has also been shown with significant reduction in the diabetes-protective effect typically associated with this haplotype [20].

A role for additional genes in the class III region in genetic susceptibility to T1D and other autoimmune disorders has been indicated by many studies (references are given in 17). For example, complement deficiency (C4A/C4B) has been associated with greater risk of T1D and younger age of onset among DR3/DR4 individuals; polymorphisms at the Major Histocompatibility Complex (MHC) class I chain-related A (*MICA*) gene have been associated with T1D susceptibility and younger age of onset; however, note that the results are heterogeneous and not consistent (reviewed in 21); and a role of the cytokine tumour necrosis factor (TNF-α) in the pathogenesis of autoimmune disorders has been reported. In all these cases, these associations may reflect LD with other causative genes, in the class III region or elsewhere, rather than a direct role of these genes. The meta-analysis of *MICA* data is particularly illustrative of this possibility [21]. Additionally, functional studies of the *BAT1* gene, which is between the *TNFA* and class I *HLA-B* gene, implicate it as a good candidate gene for T1D and other autoimmune diseases [22,23].

In this study, we have investigated single nucleotide polymorphisms (SNPs) in the class III region in Caucasian families of the T1DGC data set using the OCGM. Using a replication strategy, with the data stratified by HLA DR-DQ, a number of SNPs were identified as indicating a role of a gene or genes in the class III region in risk for T1D.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Materials and Methods

Markers

We selected 356 polymorphic bi-allelic markers mapping to the MHC class III from the first oligonucleotide pool assay (OPA1) provided to investigators contributing to the T1DGC Fine Mapping effort for analysis. This set of markers covered 1.1 megabases between positions 31 431 523 and 32 519 286 on chromosome 6.

Samples

These markers were analysed in a first round on the subset of 289 HBDI Caucasian multiplex families and in a second round on 597 additional Caucasian T1DGC families from the April T1DGC 2006 release which have been described elsewhere [24]. All study participants or their parents/surrogates gave written informed consent to participate and the study protocol was approved by the relevant Ethics Committees and Institutional Review Boards. Criteria for T1D diagnosis and further information on sample collection can be found at the T1DGC web site (http://www.t1dgc.org).

Statistical Methods

Patient and Control Genotypes—One patient genotype was analysed per family, the proband from each affected sib-pair family. Control genotype frequencies were estimated from the affected family-based controls [25,26]: the non-transmitted genotypes in trio families, similarly using the proband in affected sib-pair families (conservative bias). The subtraction method was used to avoid estimation of haplotypes: for every locus, the two alleles in the patient genotype are subtracted from the four alleles in the two parents to give the control multilocus genotype [26].

For the purposes of the current study, the discovery HBDI set contributed 289 independent cases and 289 independent controls, derived from family-based analysis. Similarly, the replication (T1DGC) set contributed 597 independent cases and 597 independent controls.

Unconditional Genotype Analysis—Unconditional association: genotype frequencies in patients vs. controls were compared using a Pearson's chi-squared test with 2 d.f.

Conditional Genotype Analysis—Under the null hypothesis, the locus we have denoted *B* (individual SNPs of the OPA1 set) is neutral with respect to disease and the *A* locus accounts for all disease risk in this genetic region. In our case, the A locus is represented by DRB1–DQA1–DQB1. The frequencies of genotypes in controls at the *A* locus are denoted f_c (A_iA_j) and for patients f_p (A_iA_j), and similarly for the two locus genotypes The expected genotype frequencies among patients are given by:

$$\exp f_{\mathrm{p}}(A_iA_jB_kB_l)=f_{\mathrm{c}}(A_iA_jB_kB_l)[\,f_{\mathrm{p}}(A_iA_j)/f_{\mathrm{c}}(A_iA_j)], \tag{1}$$

and summing over all *A* locus genotypes:

$$\exp f_{\mathrm{p}}(B_hB_k)=\sum_{i=1}^{m}\sum_{j=i}^{m}\frac{f_{\mathrm{p}}(A_iA_j)}{f_{\mathrm{c}}(A_iA_j)}f_{\mathrm{c}}(A_iA_jB_hB_k). \tag{2}$$

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Because of use of the $f_p(A_iA_j)/f_c(A_iA_j)$ ratio in the estimation of expected values, application of a standard test of homogeneity of the B genotype observed (obs) and expected (exp) numbers does not give a chi-squared distribution, in fact the distribution is exponential. Thomson and Valdes [9] proposed the following test statistic when comparing exp $f_p(A_iA_jB_kB_l)$ to the observed value $f_p(A_iA_jB_kB_l)$, for consistency we refer to the latter as obs $f_p(A_iA_jB_kB_l)$:

$$\chi^2_{2df} = \frac{Np + Nc}{2} \sum_{h,k} c_{hk}, \tag{3}$$

Where

$$c_{hk} = \frac{(\exp f_p(B_hB_k) - \text{obs } f_p(B_hB_k))^2}{4(\exp f_p(B_hB_k) + \text{obs } f_p(B_hB_k))} \tag{3}$$

and exp $f_p(B_hB_k)$ is given by eqn (2). Note that no estimates of haplotypes or LD between class II loci and class III markers are needed.

Correlation Between SNPs—LD (r^2) between pairs of significantly associated SNPs was computed using Haploview version 3.32.

Correction for Multiple Comparisons—The false discovery rate (FDR; 27) method, which controls for the expected proportion of false positives, was used to adjust for multiple comparisons.

Results

Non-stratified Analyses

As expected, many SNPs showed very strong associations with disease. These are strongest near the *HLA DR-DQ* genes, but significant effects are seen across the class III region (figure 1). These may all reflect LD with the primary HLA DR-DQ associations.

Stratified Analyses Using the Overall Conditional Genotype Method

Application of the OCGM to the 289 HBDI families found significant evidence ($p < 0.05$) from 33 SNPs of an effect additional to HLA DR-DQ (figure 2). Thirteen of these SNPs were also significant ($p < 0.05$) with OCGM analysis of the additional 597 Caucasian T1DGC families (figure 2), out of a total of 44 SNPs significant by OCGM analysis on this second set. All 13 markers significant on both the HBDI and T1DGC families, and an additional three, were significant in the overall set after adjusting for multiple testing using the FDR correction of [27] (figure 3). These 16 significantly associated SNPs map to several genes (table 1) and are not all independent of each other (figure 4).

The strongest association within class III markers was with rs2395106 that maps 5′ of the *NOTCH4* gene; this SNP was not in significant LD with any other markers (figure 4). The second strongest association was with rs707915 that maps to the *MSH5* gene; this SNP is in a block of six markers all in high LD with each other and all significantly associated with T1D after adjusting for LD with class II. A third independent association was seen with rs7762619 (5′ of the lymphotoxin–alpa gene (LTA)) but this SNP is in high LD with two other markers (rs1800750 and rs4151659; figure 4). Some of the other associations, which are not because

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

of LD with other significantly associated markers, include rs3129963, rs17421133 and rs2857595. A marker in *MICA* rs1063632 was also found to be associated with T1D although this marker is in weak LD with some of the other SNPs.

Discussion

Our preliminary analyses of the T1DGC data set provide convincing evidence for a role of additional risk genes in the class III region. These data validate the power of analyses using the OCGM. This is a very simple contingency table–based method that requires no specialized software. We note several study limitations: the joint analysis of T1DGC and HBDI families did not account for potential differences between the various collecting sites; therefore, this could theoretically introduce some bias. However, the T1DGC used stringent criteria to ensure that diagnosis and ascertainment criteria were comparable across the various regions that collected samples [24], and all individuals in the study are of European descent. Therefore, we believe results from the various regions are homogenous clinically and that by selecting only markers that are associated in both the HBDI and T1DGC-collected families, we are identifying variants that should be consistently associated with T1D in Caucasians.

Although some of the associated class III markers are in LD with each other, we identified at least six associations not because of LD within class III nor with DR-DQ. Further analyses are required to exclude those class III associations because of LD with class I or with other loci in the HLA region, such as DPB1. However, these data indicate that more than one locus within the region is modifying susceptibility to T1D. A number of the markers identified map to or near genes involved in immune regulation and autoimmune disease (e.g., *LTA, TNF, NFKBIL1, LY6G5B, MICA*) some of which have been previously implicated in *T1D* ([21–23]. The strongest association identified by our study maps close to the *NOTCH4* gene, which has been implicated in susceptibility to rheumatoid arthritis, another autoimmune disease with a strong HLA component [28,29]. This gene is a member of the Notch family of transmembrane receptors that is expressed primarily on endothelial cells. Activation of Notch in various cell systems has been shown to regulate cell fate decisions. As this is a signalling molecule, if a consistent association between this gene and T1D is confirmed and a functional link can be established, it may represent a potential therapeutic target.

Our data strongly suggest that further mapping work in the class III region can lead to more than one gene independently involved in T1D susceptibility.

Acknowledgments

Research was supported in part by NIH contract N01-AI-40076 (G. T.). This research utilizes resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418.

References

1. Thomson G, Valdes AM, Noble JA, et al. Relative predispositional effects of HLA class II DRB1-DQB1 haplotypes and genotypes on type 1 diabetes: a meta-analysis. Tissue Antigens 2007;70:110–127. [PubMed: 17610416]
2. Thomson, G.; Barcellos, LF.; Valdes, AM. Searching for additional loci in a genetic region. In: Rao, DC.; Gu, CC., editors. Genetic Dissection of Complex Traits. Academic Press; San Diego, CA, USA: 2009. p. 253-292.in press
3. Hanifi Moghaddam P, de Knijf P, Roep BO, et al. Genetic structure of IDDM1: two separate regions in the major histocompatibility complex contribute to susceptibility or protection. Diabetes 1998;47:263–269. [PubMed: 9519723]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

4. Lie BA, Todd JA, Pociot F, et al. The predisposition to type 1 diabetes linked to the human leukocyte antigen complex includes at least one non-class II gene. Am J Hum Genet 1999;64:793–800. [PubMed: 10053014]
5. Robinson WP, Thomson G, Rich SS, Barbosa J. The 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]
6. Valdes AM, Thomson G. Detecting disease-predisposing variants: the haplotype method. Am J Hum Genet 1997;60:701–716.
7. Valdes AM, McWeeney S, Thomson G. HLA class II DR-DQ amino acids and insulin-dependent diabetes mellitus: application of the haplotype method. Am J Hum Genet 1997;60:717–718. [PubMed: 9042932]
8. Koeleman BPC, Dudbridge F, Cordell HJ, Todd JA. Adaptation of the extended transmission/disequilibrium test to distinguish disease associations of multiple loci: the Conditional Extended Transmission/Disequilibrium Test. Ann Hum Genet 2000;64:207–213. [PubMed: 11246472]
9. Thomson G, Valdes AM. Conditional genotype analysis: detecting secondary disease loci in linkage disequilibrium with a primary disease locus. BMC Proc 2007;1(Suppl 1):S163. [PubMed: 18466509]
10. Thomson G. HLA DR antigens and susceptibility to insulin dependent diabetes mellitus. Am J Hum Genet 1984;36:1309–1317. [PubMed: 6334991]
11. Valdes AM, Noble JA, Génin E, Clerget-Darpoux F, Erlich HA, Thomson G. Modeling of HLA class II susceptibility to type 1 diabetes reveals an effect associated with DPB1. Genet Epidemiol 2001;21:212–223. [PubMed: 11668578]
12. 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. [PubMed: 10615959]
13. Valdes AM, Thomson G, Erlich HA, Noble JA. Association between IDDM age of onset and HLA among sib pairs. Diabetes 1999;48:1658–1661. [PubMed: 10426387]
14. 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. [PubMed: 15784469]
15. 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. [PubMed: 12121673]
16. Johansson S, Lie BA, Todd JA, et al. Evidence of at least two type 1 diabetes susceptibility genes in the HLA complex distinct from HLA-DQB1, -DQA1 and -DRB1. Genes Immun 2003;4:46–53. [PubMed: 12595901]
17. Valdes AM, Wapelhorst B, Concannon P, Erlich HA, Thomson G, Noble JA. Extended DR3-D6S273-HLA-B haplotypes are associated with increased susceptibility to type 1 diabetes in US Caucasians. Tissue Antigens 2005;65:115–119. [PubMed: 15663750]
18. Pugliese, A.; Dorman, JS.; Steenkiste, A., et al. The 13th International Histocompatibility Working Group for Type 1 Diabetes (T1D) Joint Report. In: Hansen, JA., editor. Immunobiology of the Human MHC: Proceedings of the 13th International Histocompatibility Workshop and Congress. Vol. I. Seattle, WA: IHWG Press; 2007. p. 788-796.
19. Steenkiste A, Valdes AM, Feolo M, et al. 14th International HLA and Immunogenetics Workshop: report on the HLA component of type 1 diabetes. Tissue Antigens 2007;69(Suppl 1):214–225. [PubMed: 17445204]
20. Valdes AM, Thomson G, Graham J, et al. D6S265*136 marks a DRB1*15 DQB1*0602 haplotype associated with attenuated protection from type 1 diabetes. Dia-betologia 2005;48:2540–2543.
21. Alizadeh BZ, Eerligh P, van der Slik AR, et al. MICA marks additional risk factors for type 1 diabetes on extended HLA haplotypes: an association and meta-analysis. Mol Immunol 2007;44:2806–2812. [PubMed: 17350686]
22. Wong AML, Allcock RJN, Cheong KYM, Christiansen FT, Price P. Alleles of the proximal promoter of BAT1, a putative anti-inflammatory gene adjacent to the TNF cluster, reduce transcription on a disease-associated MHC haplotype. Genes Cells 2003;8:403–412. [PubMed: 12653967]
23. Price P, Wong AML, Williamson DW, et al. Polymorphism at positions -22 and -348 in the promoter of the BAT1 gene affect transcription and the binding of nuclear factors. Hum Mol Genet 2004;13:967–974. [PubMed: 15028669]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

24. 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]

25. Thomson G. Analysis of complex human genetic traits: an ordered-notation method and new tests for mode of inheritance. Am J Hum Genet 1995;57:474–486. [PubMed: 7668274]

26. Thomson G. Mapping disease genes: family based association studies. Am J Hum Genet 1995;57:487–498. [PubMed: 7668275]

27. Benjamini Y, Hochberg T. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Statist Soc Ser B 1995;57:289–300.

28. Tazi-Ahnini R, Timms JM, Cox A, Wilson AG. Identification of novel single nucleotide polymorphisms within the NOTCH4 gene and determination of association with MHC alleles. Eur J Immunogenet 2003;30:101–105. [PubMed: 12648276]

29. Kochi Y, Yamada R, Kobayashi K, et al. Analysis of single-nucleotide polymorphisms in Japanese rheumatoid arthritis patients shows additional susceptibility markers besides the classic shared epitope susceptibility sequences. Arthritis Rheum 2004;50:63–71. [PubMed: 14730600]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

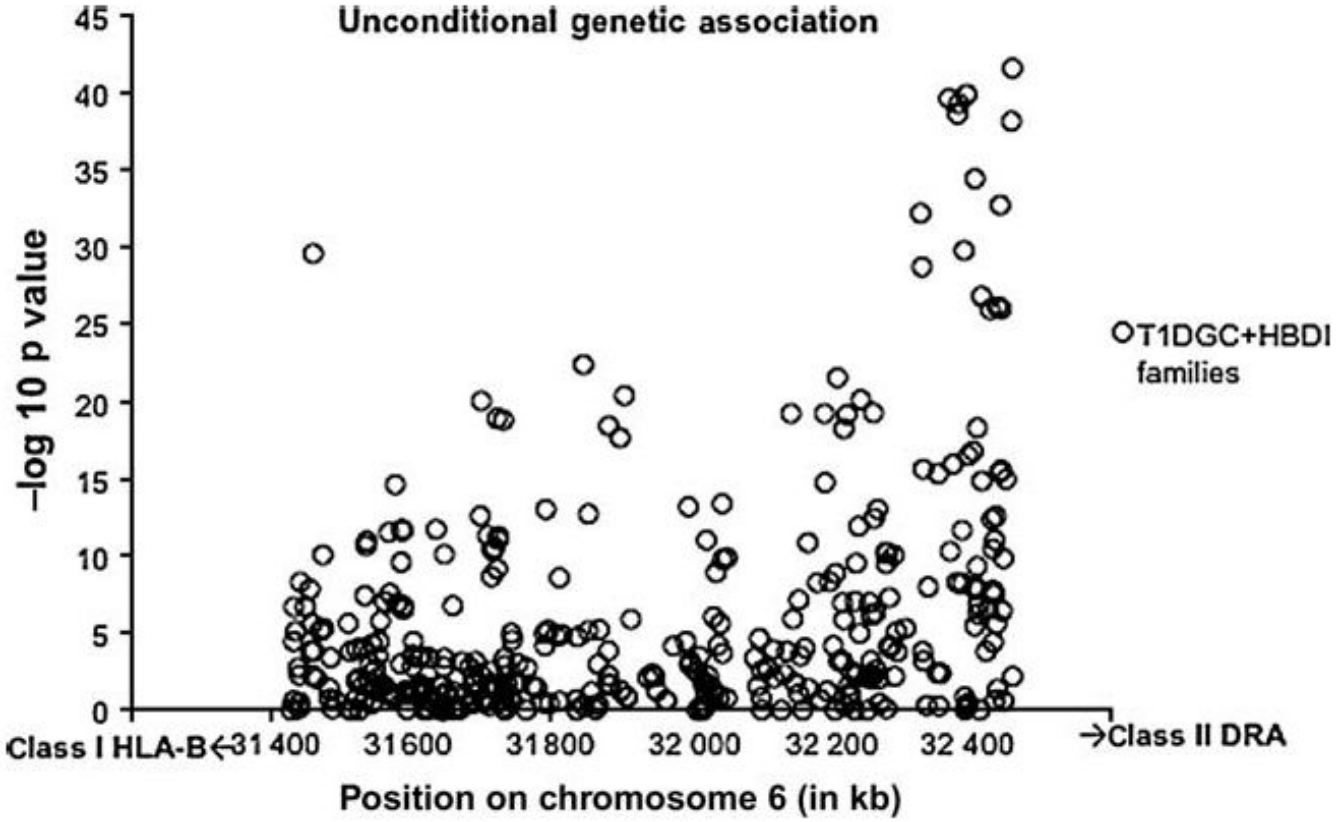

Fig. 1.

Unconditional genotype association between class III markers and type 1 diabetes in 896 Caucasian families.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

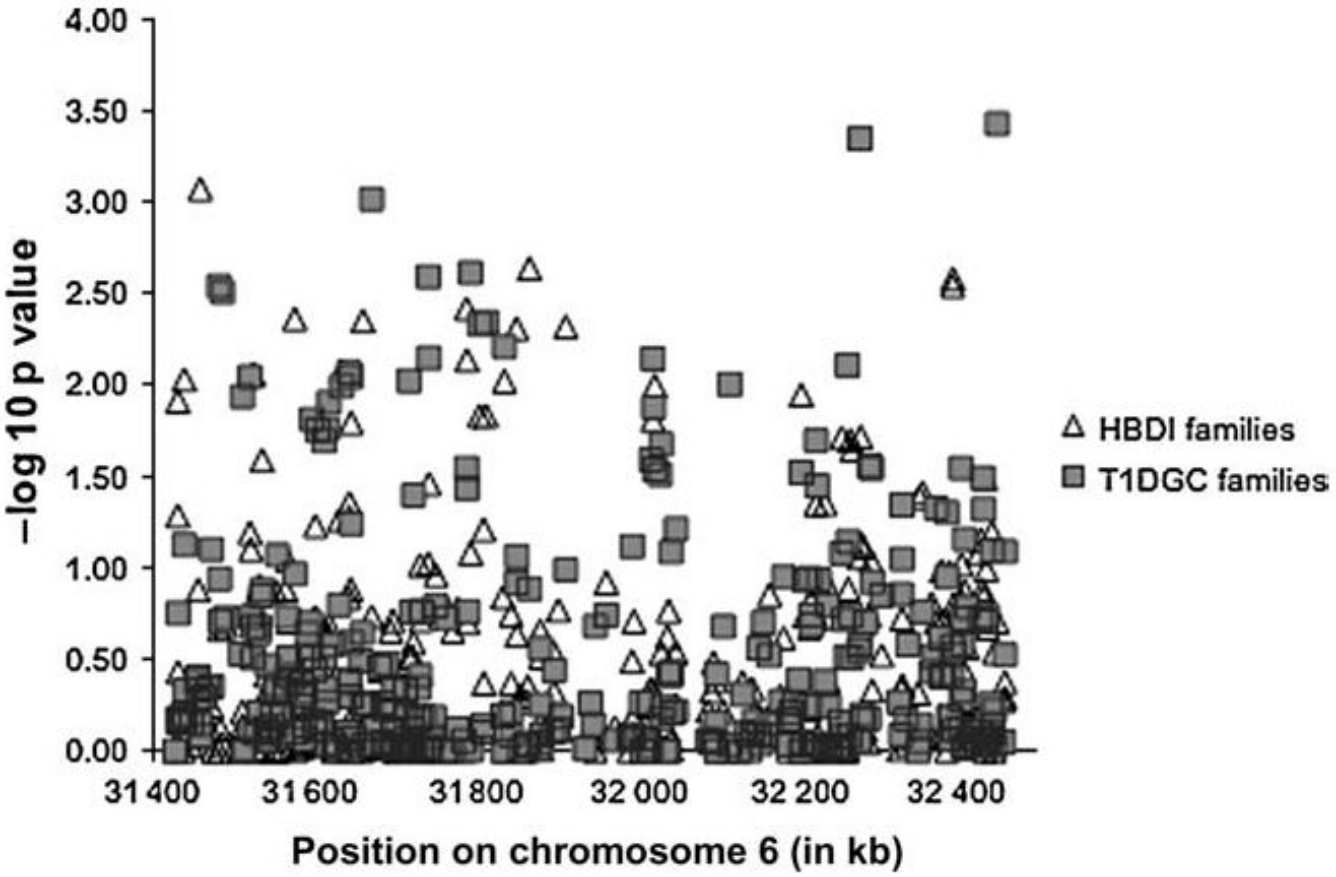

Fig. 2.
Conditional genotype association between class III markers and type 1 diabetes in two subsets of Caucasian families.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

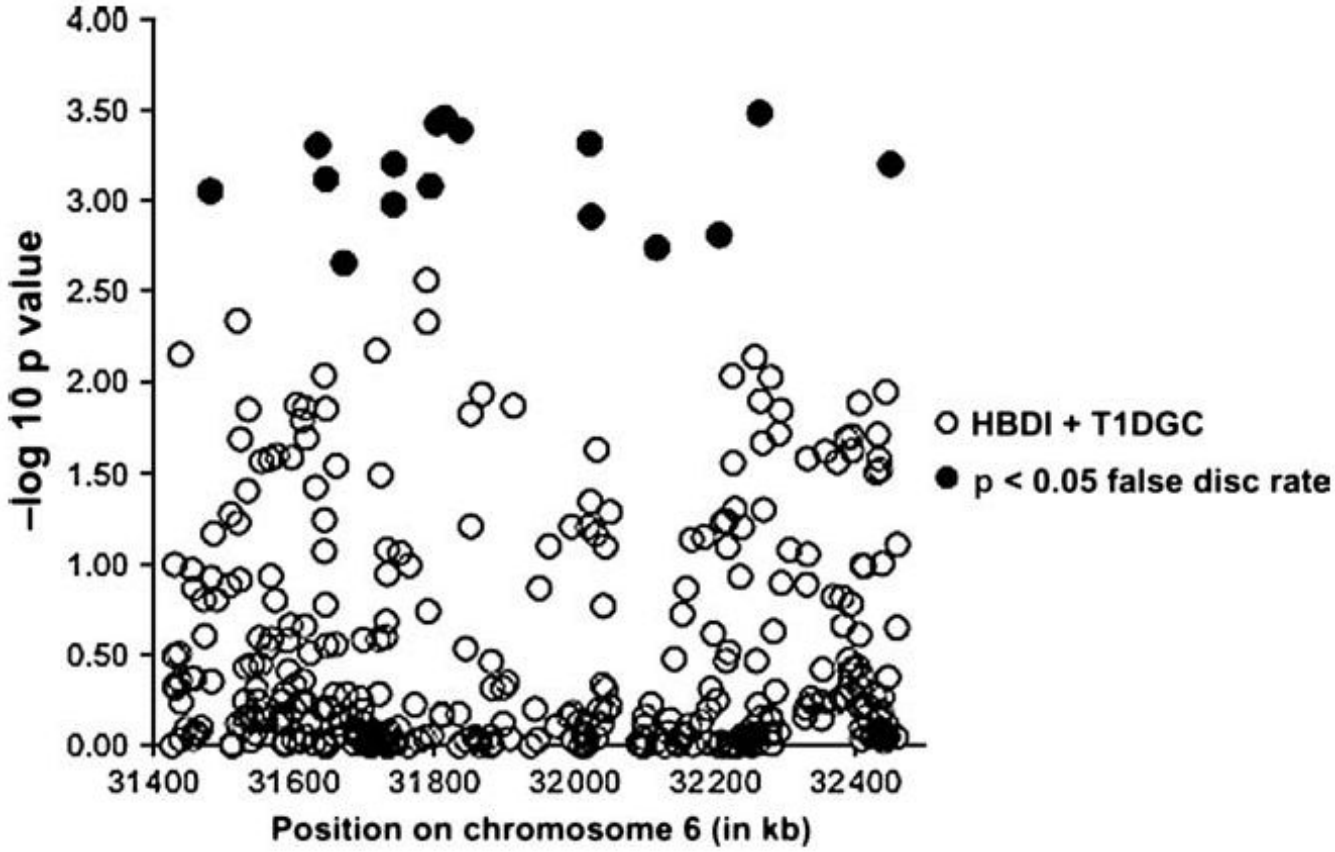

Fig. 3.

Conditional genotype association between class III markers and type 1 diabetes in 896 Caucasian families. Markers significant after correction for multiple tests are highlighted with filled symbols.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

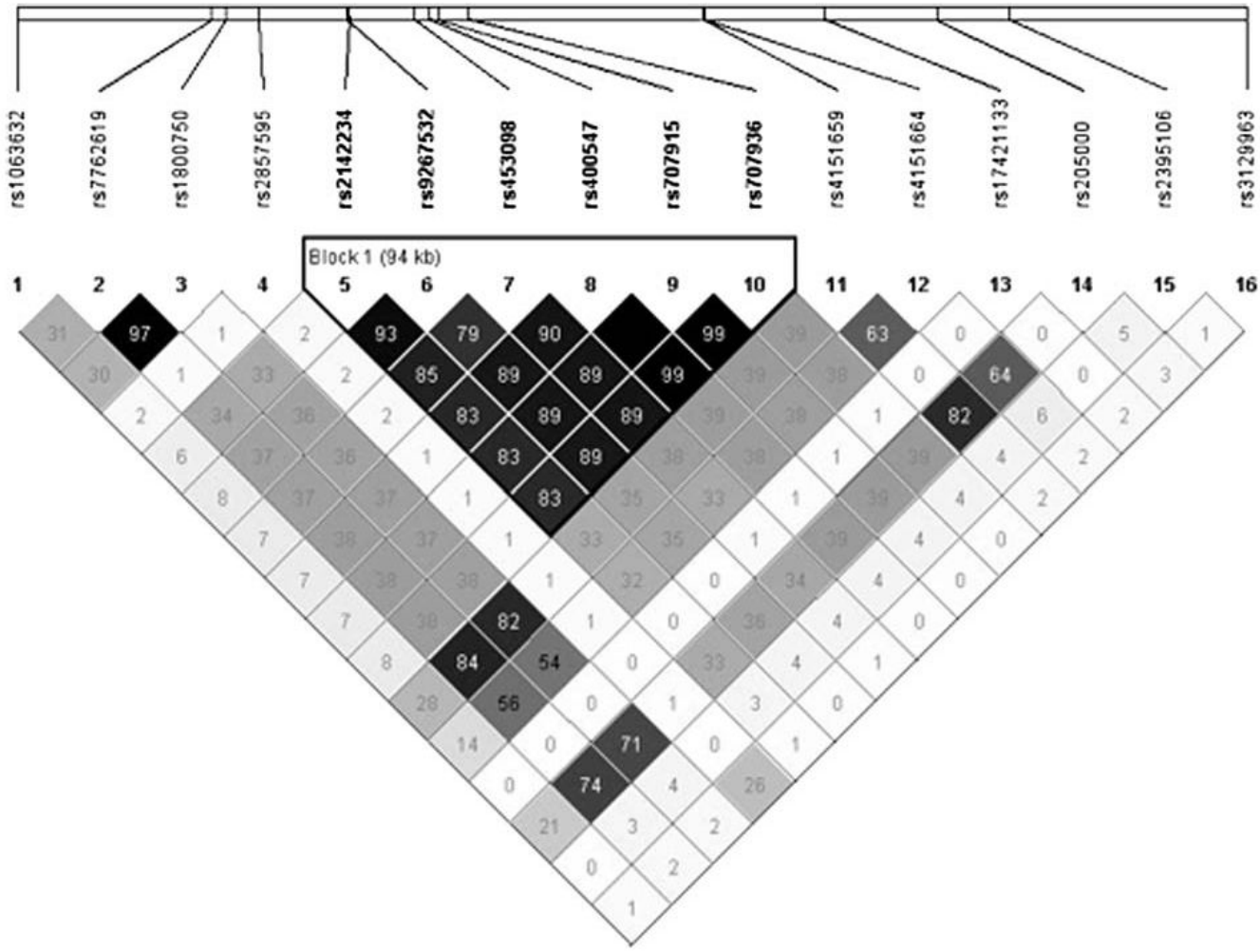

Fig. 4.

Pairwise linkage disequilibrium between class III SNPs associated with type 1 diabetes. For each pair of SNPs, the $r^2 \times 100$ is shown with darker squares representing higher LD.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

SNPs showing significant association with type 1 diabetes after adjusting for linkage disequilibrium with class II genotypes

SNP	Position on chromosome 6 (kb)	Nearest gene(s)	p value joint analysis	MAF (%)
rs1063632	31482.7	*MICA*	0.00089	10.9
rs7762619	31635.6	*3′ NFKBIL1, 5′ LTA*	0.00050	6.1
rs1800750	31647.3	*TNF*	0.00076	6.2
rs2857595	31672.8	5′ of *NCR3*	0.00219	24.7
rs2142234	31743.4	*LY6G5B*	0.00105	13.0
rs9267532	31744.3	*LY6G5B*	0.00063	12.0
rs453098	31795.9	*C6orf25*	0.00083	12.6
rs400547	31807.0	*CLIC1*	0.00036	12.6
rs707915	31815.2	*MSH5*	0.00035	12.1
rs707936	31837.9	*C6orf27*	0.00041	12.1
rs4151659	32022.7	*CFB*	0.00048	6.4
rs4151664	32025.1	*RDBP*	0.00123	8.6
rs17421133	32118.3	*TNXB*	0.00182	14.2
rs205000	32207.3	*EGFL8*	0.00155	7.1
rs2395106	32264.7	*5′ NOTCH4*	0.00033	48.7
rs3129963	32451.4	*5′ of BTNL2*	0.00063	33.5

SNP, single nucleotide polymorphisms; TNF, tumour necrosis factor; MAF, Minor allele frequency.