Published in final edited form as:
Genes Immun. 2009 January ; 10(1): 5–10. doi:10.1038/gene.2008.82.

CD226 Gly307Ser association with multiple autoimmune diseases

Jason P Hafler[1], **Lisa M Maier**[2], **Jason D Cooper**[1], **Vincent Plagnol**[1], **Anne Hinks**[3], **Matthew J Simmonds**[5], **Helen Stevens**[1], **Neil Walker**[1], **Barry Healy**[1], **Joanna M M Howson**[1], **Meeta Maisuria**[1], **Simon Duley**[1], **Gillian Coleman**[1], **Stephen C L Gough**[5], **Jane Worthington**[3], **Vijay K. Kuchroo**[2], **Linda S Wicker**[1], **John A Todd**[1], and **The International Multiple Sclerosis Genetics Consortium (IMSGC)**

[1]Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research, Department of Medical Genetics, University of Cambridge, Addenbrooke's Hospital, Cambridge, CB2 0XY, UK.

[2]Division of Molecular Immunology, Center for Neurologic Diseases, Brigham & Women's Hospital and Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA.

[3]Arthritis Research Campaign (arc)-Epidemiology Unit, Stopford Building, The University of Manchester, Manchester, M13 9PT, UK.

[5]Division of Medical Sciences, University of Birmingham, Birmingham B15 2TT, UK.

Abstract

Genome-wide association (GWA) studies provide insight into multigenic diseases through the identification of susceptibility genes and etiological pathways. In addition, identification of shared variants among autoimmune disorders provides insight into common disease pathways. We previously reported association of a nonsynonymous single nucleotide polymorphism (nsSNP) rs763361/Gly307Ser in the immune response gene *CD226* on chromosome 18q22 with type 1 diabetes (T1D) susceptibility. Here, we report efforts towards identifying the causal variant by exonic resequencing and tag SNP mapping of the 18q22 region in both T1D and multiple sclerosis (MS). In addition to the analysis of newly available samples in T1D (2,088 cases and 3,289 controls) and autoimmune thyroid disease (AITD) (821 cases and 1,920 controls), resulting in strong support for the Ser^{307} association with T1D ($P = 3.46 \times 10^{-9}$) and continued potential evidence for AITD ($P = 0.0345$), we provide convincing evidence for association of Gly307Ser with MS ($P = 4.20 \times 10^{-4}$) and some evidence for another autoimmune disease, rheumatoid arthritis (RA) ($P = 0.017$). The Ser^{307} allele of rs763361 in exon 7 of *CD226* predisposes to T1D, MS, possibly AITD and possibly RA, and based on the tag SNP analysis, could be the causal variant.

Keywords

type 1 diabetes; multiple sclerosis; rheumatoid arthritis; CD226; DNAM-1

Corresponding author: jason.hafler@cimr.cam.ac.uk.

URLs: British 1958 Birth Cohort: http:// www.b58cgene.sgul.ac.uk/; T1DBase: http://t1dbase.org (and UK mirror site, http://dil.t1dbase.org) Stata: http://www.stata.com/; R: http://www.r-project.org/; rpart: http://cran.r-project.org/; Haploview: http://www.broad.mit.edu/mpg/haploview/; gbrowse: http://www.gmod.org/; bioconductor: http://www.bioconductor.org/; dbSNP: http://www.ncbi.nlm.nih.gov/projects/SNP/

Europe PMC Funders Author Manuscripts

Introduction

Type 1 diabetes (T1D), multiple sclerosis (MS), autoimmune thyroid disease (AITD) and rheumatoid arthritis (RA) are organ specific autoimmune diseases mediated by self-reactive T cells and other cells of the adaptive and innate immune systems. T1D is characterized by inflammation of the pancreatic islets of Langerhans with destruction of insulin producing ß-cells1, while in MS and RA, there is selective white matter or joint tissue destruction, respectively2; 3. Sibling and twin studies indicate a major genetic component of the familial clustering of these common diseases4-6 with the major susceptibility loci in the HLA region7-10. Recent genome-wide association (GWA) studies have successfully identified many non-HLA single nucleotide polymorphisms (SNPs) associated with common disease susceptibility11. In addition to the identification of associated SNPs, these investigations are providing insight into genes and mechanisms shared among autoimmune diseases. Examples include *STAT4* in RA and systemic lupus erythematosus (SLE)12, and *IL2RA* in T1D13, Graves' disease14 and MS15; 16.

We recently reported a GWA study of nsSNPs in T1D that provided strong statistical evidence for association at chromosome 18q2217. In 6,021 T1D cases and 6,088 controls, Gly307Ser, located in *CD226*, showed a $P = 2.82 \times 10^{-8}$ (Odds Ratio (OR) for minor allele = 1.16, 95% confidence interval (c.i.) = 1.10-1.22) and in 2,997 parent-child trios, with a $P = 0.0281$ (Relative Risk, RR, = 1.08, 95% c.i. = 1.00-1.16). Combining the results obtained from the case-control and family studies yielded a $P = 1.38 \times 10^{-8}$.

CD226 (also known as DNAX accessory molecule 1, DNAM-1) is a 67 kDa type I membrane protein involved in the adhesion and co-stimulation of T cells18. It belongs to the immunoglobulin supergene family of receptors, containing two Ig-like domains in the extracellular region and is constitutively expressed on the majority of natural killer (NK) cells, $CD4^+$ and $CD8^+$ T cells, monocytes, platelets and a subset of B cells18. Further, in an experimental model of MS, experimental autoimmune encephalomyelitis (EAE), anti-CD226 mAb treatment delayed the onset and reduced the severity of EAE19. The Gly307Ser variant could alter expression or signalling of CD226 as it occurs in the molecule's cytoplasmic tail17. It was, therefore, of interest to explore whether Gly307Ser was the causal variant in the region and a shared risk locus for autoimmune disease.

Here we report an initial fine-mapping study of the 18q22 region in both T1D and MS by means of exonic resequencing and a tag SNP mapping approach based around Gly307Ser. We provide no evidence against the hypothesis that the nonsynonymous SNP (Gly307Ser) is the causal variant in the 18q22 region for MS and T1D. Moreover, we extended this analysis to include RA (and additional AITD samples), suggesting that Ser^{307} predisposes to a range of human autoimmune diseases.

Results

In order to increase SNP density and detect as-yet-unknown SNPs in the coding region or identify SNPs that may disrupt intron/exon splice sites present in *CD226,* we resequenced the exonic regions in the roughly 50 kb linkage disequilibrium (LD) block (exons 4, 5, 6, and 7) containing Gly307Ser and 3 kb of 3′ flanking sequence in 32 individuals chosen from the HapMap CEPH collection. This led to 7.7 kb of DNA being resequenced and the identification of 13 SNPs in three exons and the 3′ flanking sequence (exon 5 was not sequenced due to PCR failures, see material and methods). When compared with the publicly available SNPs in dbSNP build 128, two SNPs were found to be novel polymorphisms. They were located in exon 6: a nsSNP (Ala279Leu, ss102661466) and a synonymous SNP (Gln282Gln, ss102661465) with minor allele frequencies (MAFs) of

0.065 and 0.078, respectively. As these variants are functional candidates, they were genotyped in the T1D case-control collection. The single locus tests provided little evidence of an association with T1D disease susceptibility: Ala279Leu $P = 2.08 \times 10^{-3}$ (OR = 1.15; 95% c.i. 1.05-1.26) and Gln282Gln $P = 0.0298$ (OR = 0.90; 95% c.i. 0.82-0.99) (Supplementary Information Table 1). Nor did the forward logistic regression analysis adding either novel SNP to Gly307Ser provide evidence (minimum $P = 0.0542$) of an independent association with T1D susceptibility, while Gly307Ser added significantly to both SNPs (minimum $P = 8.10 \times 10^{-6}$) indicating that neither novel SNPs are independently associated with T1D susceptibility.

Further, we selected and tested a set of tag SNPs (see materials and methods) to investigate the association previously identified in the 18q22 region. Gly307Ser was still the most associated with T1D in 8,109 cases and 9,377 controls ($P = 1.32 \times 10^{-8}$; OR = 1.13, 95% c.i. 1.08-1.18) (Table 1). We conducted a forward logistic regression analysis testing the addition of each SNP to Gly307Ser and found that none added significantly. There was, therefore, no evidence for a known polymorphism (with a MAF > 0.05 and r^2 with Gly307Ser > 0.25) in the *CD226* region that showed stronger association with T1D than Gly307Ser or had an independent effect on T1D susceptibility.

We then proceeded to test Gly307Ser in a cohort of MS samples consisting of 1,275 trios, 1,194 USA cases, 595 USA controls, 993 UK cases and 9,377 UK controls (UK controls are the same as in our T1D association study). The combined *P*-value was 4.20×10^{-4} (Table 2). In order to test the hypothesis that Gly307Ser was the causal variant in MS, we tested the same set of T1D tag SNPs in an extended set of 1,318 MS trios, 1,769 MS US cases, 2,508 US controls, and 1,003 MS UK cases and used the genotyping data already available for 9,377 UK controls. Consistent with our T1D study, we obtained no evidence against the hypothesis that *CD226* Gly307Ser is the causal variant associated with MS in the 18q22 region (Table 3).

As Gly307Ser was found to be associated with both T1D and MS susceptibility, it was of interest to examine a collection of RA samples and an additional cohort of autoimmune thyroid disease cases. We tested Gly307Ser in 3,595 RA cases and 3,214 controls and obtained some evidence of association, at $P = 0.017$ (OR = 1.09; 95% c.i. 1.02-1.16) (Table 2). We obtained no evidence of heterogeneity of association between males and females ($P = 0.90$), nor between RF positive and RF negative cases ($P = 0.86$), nor between anti-CCP negative and anti-CCP positive cases ($P = 0.45$). This suggests Gly307Ser is associated with RA and not a sub-phenotype. Further, adding 821 AITD cases to our previously published data17 (N = 2,958, N = 5,431) we obtained potential evidence for association, at $P = 0.0335$ (OR = 1.08; 95% c.i. 1.00 - 1.15) (the controls are the same used in our T1D association study but matched geographically) (see Supplementary table 4 for Graves' disease and Hashimoto's disease results separately reported).

Discussion

GWA scans in common human autoimmune diseases have recently identified many loci associated with disease susceptibility. Understanding allelic heterogeneity and homogeneity among diseases provides insight into common gene function and pathways. Here, we examined the gene encoding CD226, a molecule expressed on the surface of haematopoietic cells that has independently been implicated in the pathogenesis of animal models of autoimmune diseases. Our resequencing efforts and tagging approach aimed at narrowing the association in the 18q22 region provided no evidence against the hypothesis that *CD226* Gly307Ser is the causal variant in T1D and MS. In addition, the International Multiple Sclerosis Genetics Consortium (IMSGC) extended the evidence supporting an association of

Gly307Ser with MS (see accompanying short report). In an additional sample of 3,610 MS cases, 324 controls and 1,036 trios, the IMSGC further validated Gly307Ser association with MS ($P = 5.4 \times 10^{-8}$) (Table XX in IMSGC paper, see editorial for complete analysis and overlap between studies). Taken together with our association study of the role of Gly307Ser in a collection of RA cases and additional data for Gly307Ser in AITD ($P = 0.0345$), we provide initial evidence for the *CD226* gene to be shared among at least four common human autoimmune diseases.

We note, however, that until a more complete set of polymorphisms is identified and genotyped in a large collection of cases and control subjects, we cannot exclude another variant in LD with Gly307Ser being the causal variant. Future successful resequencing of exon 5 may provide as yet undiscovered variants that will need to be assessed for disease susceptibility. In addition, the *CD226* region may harbour other, independent associations with susceptibility to disease that our tag mapping approach was not designed to identify, as has previously been shown for another T1D susceptibility locus containing the *IL2RA* gene20, although the data indicate that this is not the case for common variants in the *CD226* region.

CD226 is implicated in natural killer cell mediated cytoxicity as well as Th1 cell mediated immune response18; 19. Phosphorylation of the cytoplasmic tail of CD226 assists in co-localization with LFA-1 and cell activation21. Our genetic association data now justify studies of the functional consequences of the Gly307Ser variant in adaptive and innate immune responses. We have previously hypothesized17 that the SNP could disrupt a splice site enhancer, or silencer, thereby altering RNA splicing, as has been demonstrated for other immune related genes (human *CD45* and mouse *Ctla4*)22, resulting in either a putative *CD226* isoform acting as a non-functional (non-signalling) protein, or with a novel function. Alternatively, this amino acid substitution could alter the signalling cascade by affecting the two known phosphorylation sites at positions 322 and 32921; 23, which share a critical role in CD226 and the immune response.

Material and methods

Subjects

All case and control subjects were of self-reported white ethnicity and were enrolled under study protocols approved by the Institutional Review board of each institution that contributed. Written informed consent was obtained from the participants or their guardians.

Type 1 diabetes collection

T1D cases were recruited as part of the Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation Laboratory's British case collection (Genetic Resource Investigating Diabetes)17. Control samples were obtained from the British 1958 Birth Cohort (B58C), and WTCCC Blood Service controls11. Cases and controls were chosen to be matched geographically.

Multiple sclerosis collection

Healthy adult control subjects were recruited through the Brigham and Women's Hospital, the University of Cambridge and the University of California at San Francisco, as previously described15. All were unrelated individuals having no history of chronic inflammatory disease. MS cases were collected as described in our recent investigation of patients with MS15. Subjects with MS all meet McDonald criteria for MS24.

Rheumatoid arthritis collection

DNA was from UK RA patients, over 18 years old and satisfying the American College of Rheumatology criteria for RA was available from six centres in the UK and five of these centres provided controls as described in our recent investigation25. The samples collected from 6 centres in the UK raise the possibility that the results were affected by population substructure and heterogeneity. A stratified analysis by center revealed that the association of Gly307Ser with rheumatoid arthritis was independently observed in 4 of the 5 centers tested (one center had no controls, and therefore association could not be statistically tested for this centre). No heterogeneity was detected among the samples from the different centers and combined evidence from the different centers (using a Cochran-Mantel-Haenszel test) attained a significance level virtually the same as that from the combined samples.

Autoimmune thyroid disease collection

As previously reported17 as part of the autoimmune thyroid disease (AITD) UK National Collection, 2,958 unrelated individuals were recruited including 2,295 with Graves' disease and 663 with Hashimoto's Thyroiditis. Cases and controls were chosen to be matched geographically.

Sequencing

Polymorphisms were identified by resequencing 32 Centre d'Etude du Polymorphisme Humain (CEPH) DNA samples, which are the same samples used by HapMap (www.hapmap.org). Sequencing was performed using Applied Biosystems' BigDye chemistry (version 3.1) and the sequences resolved using an ABI 3700 Genetic Analyzer. Analyses of the sequence traces were performed using the Staden package, and traces were scored independently by a second operator by hand. Annotations are available from T1DBase (see URLs), together with sequence and polymorphism data at the T1DBase PosterPages (see URL). Primer sequences are available upon request. Due to problems with design of primers to amplify exon 5, this exon could not be successfully sequenced and any as yet undiscovered variants that may reside in this exon are not part of our association analysis.

Genotyping

SNPs were genotyped using the iPLEX™ Sequenom MassARRAY® platform, or TaqMan® (Applied Biosystems) in accordance with the manufacturer's instructions. Cases and controls were genotyped and data scored twice to minimize error, with the second operator being unaware of case-control status or family structure. None of the SNPs significantly deviated from Hardy-Weinberg disequilibrium in controls and unaffected parents ($P>0.05$) (except for rs17208112, see tag SNP section).

Statistical analyses

All statistical analyses were performed in the Stata or R statistical packages (see URLs).

Logistic regression analyses

The case-control data were analysed using logistic regression models stratified by 12 geographical regions across England, Scotland and Wales to minimise loss of power due to geography26. When analysing a single SNP, we performed a one degree of freedom (1 d.f.) likelihood ratio test to determine whether a 1 d.f. multiplicative allelic effects model or a 2 d.f. genotypic effects model better fit the data27.

We used forward logistic regression to assess the evidence against the most significant SNP being the sole associated variant in the region (in other words, whether this SNP alone was

sufficient to model the association). For the purposes of this analysis, we did not assume any specific mode of inheritance for the most associated SNP (A>a) or for any additional SNP with significant independent effects on T1D, so genotype risks of A/A and A/a were modeled relative to the a/a genotype. We then used a 1-d.f. test for adding each of the remaining SNPs to the model by assuming multiplicative allelic effects for the additional SNPs.

Tag SNP selection

To delimit the disease-associated region and select an informative set of tags, we analyzed the LD (using r^2 and D′28) structure of *CD226* in DNA samples obtained from 32 individuals from the CEPH collection genotyped by the International HapMap project (www.hapmap.org)29. The tagging strategy involved the selection of SNPs with minor allele frequencies (MAFs) > 0.05 and r^2 values > 0.25 with Gly307Ser. Available data allowed for the analysis of 205 SNPs within *CD226*. Of the 135 SNPs in the LD block, 43 had MAFs > 0.05 and r^2 > 0.25 with Gly307Ser. Eleven SNPs were sufficient to pairwise tag this LD block (r^2 > 0.80) as determined in HapMap (release 21). If an association were not observed in 5,500 cases and 5,500 controls, the SNP would not be genotyped in additional cases.

rs17208112, a singleton SNP tagging itself, failed quality control tests in both our T1D and UK MS cohorts due to an adjacent SNP disrupting the binding specificity of the probe and was hence removed from the data set.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by the Juvenile Diabetes Research Foundation International and the Wellcome Trust. We gratefully acknowledge the participation of all the patients and control subjects. We acknowledge use of the DNA from the British 1958 Birth Cohort collection, funded by the Medical Research Council and Wellcome Trust. We also thank The Avon Longitudinal Study of Parents and Children laboratory in Bristol and the British 1958 Birth Cohort team, including S. Ring, R. Jones, M. Pembrey, W. McArdle, D. Strachan and P. Burton for preparing and providing the control DNA samples. We thank Cristin Aubin at the Broad Institute. We thank Helen Schuilenburg and Nigel Ovington for data support as well as Oliver Burren for bioinformatics support. DNA samples were prepared by P. Clarke, J. Denesha, D. Harrison, S. Hawkins, M. Himsworth, T. Mistry, N. Taylor, N. Ubani. This study makes use of data generated by the Wellcome Trust Case Control Consortium. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk.

Bibliography

1. Haller MJ, Atkinson MA, Schatz D. Type 1 diabetes mellitus: etiology, presentation, and management. Pediatr Clin North Am. 2005; 52(6):1553–1578. [PubMed: 16301083]
2. Prineas JW, Wright RG. Macrophages, lymphocytes, and plasma cells in the perivascular compartment in chronic multiple sclerosis. Lab Invest. 1978; 38(4):409–421. [PubMed: 205724]
3. Stastny P. Mixed lymphocyte cultures in rheumatoid arthritis. J Clin Invest. 1976; 57(5):1148–1157. [PubMed: 1262462]
4. Doolittle TH, Myers RH, Lehrich JR, Birnbaum G, Sheremata W, Franklin GM, et al. Multiple sclerosis sibling pairs: clustered onset and familial predisposition. Neurology. 1990; 40(10):1546–1552. [PubMed: 2215946]
5. Silman AJ, MacGregor AJ, Thomson W, Holligan S, Carthy D, Farhan A, et al. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. Br J Rheumatol. 1993; 32(10):903–907. [PubMed: 8402000]

6. Hyttinen V, Kaprio J, Kinnunen L, Koskenvuo M, Tuomilehto J. Genetic liability of type 1 diabetes and the onset age among 22,650 young Finnish twin pairs: a nationwide follow-up study. Diabetes. 2003; 52(4):1052–1055. [PubMed: 12663480]
7. Todd JA, Bell JI, McDevitt HO. HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature. 1987; 329(6140):599–604. [PubMed: 3309680]
8. Stastny P. Association of the B-cell alloantigen DRw4 with rheumatoid arthritis. N Engl J Med. 1978; 298(16):869–871. [PubMed: 147420]
9. Jersild C, Svejgaard A, Fog T. HL-A antigens and multiple sclerosis. Lancet. 1972; 1(7762):1240–1241. [PubMed: 4113225]
10. Simmonds MJ, Howson JM, Heward JM, Carr-Smith J, Franklyn JA, Todd JA, et al. A novel and major association of HLA-C in Graves' disease that eclipses the classical HLA-DRB1 effect. Hum Mol Genet. 2007; 16(18):2149–2153. [PubMed: 17597093]
11. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007; 447(7145):661–678. [PubMed: 17554300]
12. Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med. 2007; 357(10):977–986. [PubMed: 17804842]
13. Vella A, Cooper JD, Lowe CE, Walker N, Nutland S, Widmer B, et al. Localization of a type 1 diabetes locus in the IL2RA/CD25 region by use of tag single-nucleotide polymorphisms. Am J Hum Genet. 2005; 76(5):773–779. [PubMed: 15776395]
14. Brand OJ, Lowe CE, Heward JM, Franklyn JA, Cooper JD, Todd JA, et al. Association of the interleukin-2 receptor alpha (IL-2Ralpha)/CD25 gene region with Graves' disease using a multilocus test and tag SNPs. Clin Endocrinol (Oxf). 2007; 66(4):508–512. [PubMed: 17371467]
15. Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL, et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med. 2007; 357(9):851–862. [PubMed: 17660530]
16. Weber F, Fontaine B, Cournu-Rebeix I, Kroner A, Knop M, Lutz S, et al. IL2RA and IL7RA genes confer susceptibility for multiple sclerosis in two independent European populations. Genes Immun. 2008; 9(3):259–263. [PubMed: 18354419]
17. Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet. 2007; 39(7):857–864. [PubMed: 17554260]
18. Shibuya A, Campbell D, Hannum C, Yssel H, Franz-Bacon K, McClanahan T, et al. DNAM-1, a novel adhesion molecule involved in the cytolytic function of T lymphocytes. Immunity. 1996; 4(6):573–581. [PubMed: 8673704]
19. Dardalhon V, Schubart AS, Reddy J, Meyers JH, Monney L, Sabatos CA, et al. CD226 is specifically expressed on the surface of Th1 cells and regulates their expansion and effector functions. J Immunol. 2005; 175(3):1558–1565. [PubMed: 16034094]
20. Lowe CE, Cooper JD, Brusko T, Walker NM, Smyth DJ, Bailey R, et al. Large-scale genetic fine mapping and genotype-phenotype associations implicate polymorphism in the IL2RA region in type 1 diabetes. Nat Genet. 2007
21. Shibuya K, Lanier LL, Phillips JH, Ochs HD, Shimizu K, Nakayama E, et al. Physical and functional association of LFA-1 with DNAM-1 adhesion molecule. Immunity. 1999; 11(5):615–623. [PubMed: 10591186]
22. Ueda H, Howson JM, Esposito L, Heward J, Snook H, Chamberlain G, et al. Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. Nature. 2003; 423(6939):506–511. [PubMed: 12724780]
23. Shirakawa J, Shibuya K, Shibuya A. Requirement of the serine at residue 329 for lipid raft recruitment of DNAM-1 (CD226). Int Immunol. 2005; 17(3):217–223. [PubMed: 15684041]
24. McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol. 2001; 50(1):121–127. [PubMed: 11456302]
25. Thomson W, Barton A, Ke X, Eyre S, Hinks A, Bowes J, et al. Rheumatoid arthritis association at 6q23. Nat Genet. 2007; 39(12):1431–1433. [PubMed: 17982455]

26. Clayton DG, Walker NM, Smyth DJ, Pask R, Cooper JD, Maier LM, et al. Population structure, differential bias and genomic control in a large-scale, case-control association study. Nat Genet. 2005; 37(11):1243–1246. [PubMed: 16228001]

27. Cordell HJ, Clayton DG. 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. 2002; 70(1):124–141. [PubMed: 11719900]

28. Lewontin RC. The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. Genetics. 1964; 49(1):49–67. [PubMed: 17248194]

29. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007; 449(7164):851–861. [PubMed: 17943122]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Table 1

Analysis of *CD226* tag SNPs in type 1 diabetes

rs number	MAF in controls	*P* (1-d.f. test)	OR (95% c.i.)	r^2 (with rs763361)	D′ (with rs763361)	*N* (cases)	*N* (controls)
rs763361	0.47	1.32×10^{-8}	1.13 (1.08 - 1.18)	-	-	8,109	9,377
rs1009847	0.28	6.76×10^{-4}	1.09 (1.04 - 1.15)	0.30	0.84	7,187	7,597
rs1124980	0.39	3.04×10^{-5}	1.12 (1.06 - 1.18)	0.58	0.90	6,010	5,495
rs11661553	0.43	3.25×10^{-6}	0.88 (0.84 - 0.93)	0.63	0.97	6,150	5,577
rs12604328	0.26	4.58×10^{-5}	1.13 (1.07 - 1.20)	0.37	0.96	6,082	5,596
rs17208329	0.24	0.0405	1.06 (1.00 - 1.12)	0.14	0.64	7,219	7,664
rs17842596	0.39	4.43×10^{-4}	1.09 (1.04 - 1.14)	0.29	0.63	7,339	7,608
rs1788114	0.43	1.71×10^{-6}	1.12 (1.07 - 1.18)	0.47	0.75	7,251	7,379
rs1788234	0.40	3.72×10^{-3}	1.07 (1.02 - 1.12)	0.43	0.77	7,217	7,619
rs4891781	0.40	5.23×10^{-4}	1.10 (1.04 - 1.16)	0.74	0.98	6,148	5,568

A multiplicative allelic effects model was assumed (see materials and methods), as it was not significantly different from the full genotype model. OR is reported for the minor allele MAF = minor allele frequency, OR = odds ratio, 95% c.i. = 95% confidence interval, linkage disequilibrium measures: r^2 = correlation coefficient, D′ = Lewontin's D′28, *N* = number. rs17208112, a singleton tagging itself, failed quality control checks and was therefore excluded in the analysis

Supplementary table 2 contains analysis of the tag SNPs restricted to individuals with genotype information available for all SNPs

Europe PMC Funders Author Manuscripts

Table 2

CD226 Ser307Gly (rs763361) in type 1 diabetes, multiple sclerosis, rheumatoid arthritis and Graves' disease

Disease	Cohort	MAF in controls	*P* (1-d.f. test)	OR (95% c.i.)	*P* combined	*N* (trios)	*N* (cases)	*N* (controls)
Type 1 diabetes[I]	UK (case/control)	0.47	1.32×10^{-8}	1.13 (1.08 - 1.18)	3.46×10^{-9}	-	8,109	9,377
	UK (trios)	-	0.0281	1.08 (1.00 - 1.16)[II]		3,093	-	-
Multiple sclerosis	USA & UK (trios)	-	0.0587	1.12 (1.04 - 1.19)[II]	4.20×10^{-4}	1,275	-	-
	UK (case/control)	0.47	0.0184	1.12 (1.02 - 1.23)		-	993	9,377
	USA (case/control)	0.46	0.0666	1.14 (0.99 - 1.31)		-	1,194	595
Rheumatoid arthritis	UK (case/control)	0.46	0.017	1.09 (1.02 - 1.16)	-		3,595	3,214
Graves' disease[I]	UK (case/control)	0.47	0.0335	1.08 (1.00-1.15)	-		2,958	5,431

[I] T1D and Graves' disease results previously reported. The T1D dataset originally included 6,021 cases and 6,088 controls 2 and the Graves' disease originally included 2,137 cases and 3,511 controls. The additional autoimmune thyroid disease cases include 158 Graves' disease and 663 Hashimoto's. 25 Graves' disease cases were excluded from our analysis because they also had T1D. See supplementary table 4 for separate analysis of Graves' disease and Hashimoto's disease samples

[II] Relative risk (RR)

Europe PMC Funders Author Manuscripts

Table 3

Analysis of *CD226* tag SNPs in multiple sclerosis

	Cases and controls										Families		Combined *P*
	USA					UK					Parent child trios		
rs number	MAF in controls	*P* (1-d.f. test)	OR (95% c.i.)	*N* (cases)	*N* (controls)	MAF in controls	*P* (1-d.f. test)	OR (95% c.i.)	*N* (cases)	*N* (controls)	*P* TDT	RR (95% c.i.)	
rs763361[I]	0.46	0.0666	1.14 (0.99 – 1.31)	1,194	595	0.47	0.0185	1.12 (1.02 – 1.23)	993	9,377	0.059	1.12 (1.04 - 1.19)	4.54×10^{-4}
rs1009847	0.28	0.689	1.02 (0.93 - 1.12)	1,755	2,438	0.28	0.229	1.07 (0.96 - 1.19)	963	7,187	0.30	1.07 (0.94 - 1.19)	0.089
rs1124980	0.41	0.186	1.06 (0.97 - 1.16)	1,757	2,435	0.39	0.0862	1.09 (0.99 - 1.20)	962	6,010	0.59	1.03 (0.92 - 1.15	0.00354
rs11661553	0.41	0.152	0.94 (0.86 - 1.02)	1,754	2,434	0.43	0.276	0.95 (0.86 - 1.04)	960	6,150	0.58	0.97 (0.86 - 1.08)	0.00453
rs12604328	0.27	0.258	1.06 (0.96 - 1.16)	1,752	2,433	0.26	0.294	1.06 (0.95 - 1.18)	958	6,082	0.41	1.06 (0.93 - 1.19)	0.0193
rs17208329	0.23	0.296	0.95 (0.85 - 1.05)	1,754	2,430	0.24	0.690	1.02 (0.92 - 1.14)	957	7,219	0.84	1.01 (0.89 - 1.14)	0.387
rs17842596	0.40	0.238	1.05 (0.97 - 1.15)	1,753	2,427	0.39	0.482	1.04 (0.94 - 1.14)	954	7,339	0.72	0.98 (0.87 - 1.10)	0.102
rs1788114	0.44	0.0360	1.10 (1.01 - 1.20)	1,754	2,430	0.43	0.100	1.08 (0.98 - 1.19)	961	7,251	0.86	1.01 (0.90 - 1.13)	3.91×10^{-4}
rs1788234	0.41	0.187	1.06 (0.97 - 1.16)	1,755	2,431	0.40	0.177	1.07 (0.97 - 1.18)	960	7,217	0.60	1.03 (0.92 – 1.15)	0.00802
rs4891781	0.41	0.469	1.03 (0.95 - 1.13)	1,753	2,429	0.40	0.167	1.07 (0.97 - 1.18)	956	6,148	0.10	1.10 (0.98 - 1.22)	0.015

A multiplicative allelic effects model was assumed (see methods), as it was not significantly different from the full genotype effects model. OR is reported for the minor allele MAF = minor allele frequency, *P* = *P* value, OR = odds ratio, 95% c.i. = 95% confidence interval, *N* = number, TDT = transmission-disequilibrium test.

rs17208112, a SNP only tagging itself, failed quality control checks and was therefore excluded in the analysis.

[I] We note that our original genotyping of rs763361 in USA samples allowed for ~558 fewer cases and ~1,834 fewer controls as compared to our tag SNPs. Supplementary table 3 contains an analysis of subjects with genotype information available for all SNPs: in the USA restricted data set rs1788114 is less significant (*P* = 0.22) as compared to rs763361 (*P* = 0.033).