Human Molecular Genetics, 2012, Vol. 21, No. 23 ***5202–5208***
doi:10.1093/hmg/dds357
Advance Access published on August 24, 2012

Seven newly identified loci for autoimmune thyroid disease

Jason D. Cooper[1,*,†], Matthew J. Simmonds[2,†], Neil M. Walker[1], Oliver Burren[1], Oliver J. Brand[2], Hui Guo[1], Chris Wallace[1], Helen Stevens[1], Gillian Coleman[1], Wellcome Trust Case Control Consortium[‡], Jayne A. Franklyn[3], John A. Todd[1,¶] and Stephen C.L. Gough[2,4,¶]

[1]Department of Medical Genetics, Cambridge Institute for Medical Research, Juvenile Diabetes Research Foundation/ Wellcome Trust Diabetes and Inflammation Laboratory, NIHR Cambridge Biomedical Research Centre, University of Cambridge, Addenbrooke's Hospital, Cambridge, [2]Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Churchill Hospital, Oxford, [3]School of Clinical and Experimental Medicine, Institute of Biomedical Research, University of Birmingham, Birmingham, and [4]NIHR Oxford Biomedical Research Centre, Churchill Hospital, Oxford, UK

Received May 31, 2012; Revised August 6, 2012; Accepted August 20, 2012

Autoimmune thyroid disease (AITD), including Graves' disease (GD) and Hashimoto's thyroiditis (HT), is one of the most common of the immune-mediated diseases. To further investigate the genetic determinants of AITD, we conducted an association study using a custom-made single-nucleotide polymorphism (SNP) array, the ImmunoChip. The SNP array contains all known and genotype-able SNPs across 186 distinct susceptibility loci associated with one or more immune-mediated diseases. After stringent quality control, we analysed 103 875 common SNPs (minor allele frequency >0.05) in 2285 GD and 462 HT patients and 9364 controls. We found evidence for seven new AITD risk loci ($P < 1.12 \times 10^{-6}$; a permutation test derived significance threshold), five at locations previously associated and two at locations awaiting confirmation, with other immune-mediated diseases.

INTRODUCTION

Genome-wide association (GWA) studies have had a dramatic impact on susceptibility locus discovery and in addition, highlighted and extended the previously observed (1,2) commonality of loci between immune-mediated diseases, which has ultimately resulted in the ImmunoChip project. As described previously (3), the ImmunoChip project is a collaboration between 12 immune-mediated disease groups [autoimmune thyroid disease (AITD), ankylosing spondylitis (AS), celiac disease (CeD), Crohn's disease (CD), IgA deficiency, multiple sclerosis (MS), primary biliary cirrhosis (PBC), psoriasis (PS), rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), type 1 diabetes (T1D) and ulcerative colitis (UC)] and the Wellcome Trust Case Control Consortium (WTCCC). The main objective of the project was to provide cost-effective genotyping of all known and genotype-able single-nucleotide polymorphisms (SNPs) across confirmed and distinct susceptibility loci from the 12 disease groups (www.ImmunoBase.org). The result was the ImmunoChip, a custom Illumina 200 K Infinium high-density array containing 186 distinct susceptibility loci (3), mainly derived from GWA studies, which have been convincingly associated (required $P < 5 \times 10^{-8}$) with one or more immune-mediated diseases. In addition, the ImmunoChip contained wildcard SNPs provided by the 12 disease groups and SNPs for replication found in non-immunological disease GWA studies conducted by the WTCCC (www.wtccc.org.uk/ccc2/wtccc2_studies.shtml).

The AITDs are caused by an autoimmune attack in which antibodies directed against components of the thyroid gland

*To whom correspondence should be addressed. Tel: +44 1223762107; Fax: +44 1223762102; Email jason.cooper@cimr.cam.ac.uk
†Joint first authors.
‡A full list of the Wellcome Trust Case Control Consortium investigators involved in this study are included in the Supplementary Information.
¶Joint senior authors.

© *The Author 2012. Published by Oxford University Press.*
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from http://hmg.oxfordjournals.org/ at Istituto Dermopatico Immacolata on November 6, 2012

lead to either thyroid hormone excess and clinical features of hyperthyroidism as seen in people with Graves' disease (GD) or thyroid gland damage, reduced thyroid hormone production and clinical features of hypothyroidism as seen in people with Hashimoto's thyroiditis (HT). AITD shows a strong female preponderance with estimated incidence rates of GD in females as 80 per 100 000 person-years and in males as 8 per 100 000 person-years, and of HT in females as 350 per 100 000 person-years and in males as 8 per 100 000 person-years (4).

Although many regions of the genome have been reported as AITD susceptibility loci, convincing evidence for association has been limited to four regions: *PTPN22*/1q13.2, *CTLA4*/2q33.2, HLA/6p21, and *TSHR*/14q31.1 (5). These susceptibility loci all fall within the antigen presentation/T-cell receptor signalling pathway. In addition, there is mounting evidence for a susceptibility locus in *FCRL3*/1q23.1 (5–7), *SCGB3A2*/5q32 (8,9) and *IL2RA*/10p15.1 (10,11), and a recent GWA and replication study in subjects recruited from the Chinese Han population identified two novel loci at 4p14 and *RNASET2-FGFR1OP-CCR6*/6q27 (5).

Here, we used the ImmunoChip to further investigate the genetic architecture of AITD in 2374 GD and 474 HT patients, and 9953 controls.

RESULTS

After consideration of the number of AITD patients available, we focused on susceptibility locus discovery using common SNPs [minor allele frequency (MAF) ≥ 0.05] outside of the HLA, which reduced the number of tests performed and consequently, allowed us to use a less stringent significance threshold. We adopted a permutation approach to correct the 0.05 level of significance for multiple testing and for correlation between SNP genotypes, and obtained a threshold $P = 1.12 \times 10^{-6}$ (see Materials and Methods). After stringent quality control (see Materials and Methods), we analysed 103 875 common SNP genotypes in 2285 GD and 462 HT patients, and 9364 controls. As the vast majority of our AITD patients have been previously analysed in either discovery or confirmation studies of the reported AITD susceptibility loci, we initially used the AITD susceptibility loci on the ImmunoChip as positive controls and as expected, found evidence of association ($P < 1.12 \times 10^{-6}$ and good genotype signal intensity clusters; Supplementary Material, Table S1) for *PTPN22* ($P = 9.7 \times 10^{-23}$), *CTLA4* ($P = 2.1 \times 10^{-23}$), *IL2RA* ($P = 2.7 \times 10^{-7}$) and *TSHR* ($P = 1.3 \times 10^{-38}$ in GD patients/controls). We note that the AITD risk locus *SCGB3A2* was not on the ImmunoChip, but had previously been associated in our AITD patients ($P = 0.007$; (8)). In addition, we found some evidence of association for *FCRL3* ($P = 1.1 \times 10^{-5}$ in GD patients/controls; Supplementary Material, Table S1) and found further evidence of association for the recently reported *RNASET2–FGFR1OP–CCR6*/6q27 locus (5) (imm_6_167 338101 $P = 1.6 \times 10^{-7}$; r^2 between imm_6_167338101 and rs9355610 (5) in controls = 0.72; Table 1; Supplementary Material, Fig. S1) in subjects with European ancestry. The recently reported 4p14 locus (5) was not on the ImmunoChip.

We found evidence of association between AITD and seven genomic locations ($P < 1.12 \times 10^{-6}$; Table 1). Five of these new AITD risk loci were at locations previously associated with at least one other immune-mediated disease (Table 1; Supplementary Material, Figs S2–S6); three of which included the most AITD-associated SNPs being located within MMEL1/1p36.32, LPP/3q28 and BACH2/6q15 genes. However, only *BACH2* has a known autoimmune function as a regulator of nucleic acid-triggered antiviral responses in human cells (12). The remaining two new AITD risk loci were unconfirmed risk loci: rs1534422 at 2p25.1 is an unconfirmed T1D risk locus (T1D $P = 2.1 \times 10^{-6}$; (13); Supplementary Material, Fig. S7); and rs4409785 at 11q21 is an unconfirmed MS risk locus (MS $P = 6.3 \times 10^{-7}$; (14); Supplementary Material, Fig. S8).

We tested for heterogeneity in disease association before combining the GD and HT patients (Table 1 and Supplementary Material, Table S1). After allowing for multiple testing, only the disease association of the *TSHR*/14q31.1 locus was significantly different between the two patient groups. This difference between GD and HT patients was expected as the *TSHR* represents the primary autoantigen in hyperthyroidism of GD (15). As we focused on susceptibility loci discovery using common SNPs (MAF > 0.05) outside of the HLA to reduce the number of tests performed, which was necessary given the number of GD and HT patients available in this study, we did not proceed to test whether the most disease-associated SNP alone was sufficient to model the association signal in AITD risk loci with dense SNP coverage.

We examined the overlap between AITD and the other 11 immune-mediated diseases that provided susceptibility loci to the ImmunoChip in turn, by testing whether the pattern of AITD association across immune-mediated disease regions differed according to whether those regions had previously been associated with each specific immune-mediated disease or not (Table 2; (16)). We restricted the analysis to non-HLA SNPs from densely mapped regions associated with at least one immune-mediated disease. We found that SNPs within RA-associated regions were significantly ($P = 6.56 \times 10^{-5}$) more likely to be associated with AITD than SNPs within regions not associated with RA. In addition, there was marginal evidence of an overlap between AITD and T1D-associated regions ($P = 0.0327$). In contrast, we found that SNPs within regions associated with AS, UC and PS showed significantly less evidence of association with AITD than SNPs within regions known to be associated with each disease ($P = 3.27 \times 10^{-4}$, 1.69×10^{-3} and 4.08×10^{-3}, respectively). There was also marginal evidence of an overlap between AITD and regions not associated with CD ($P = 0.0250$). We note that an important caveat about interpreting these test results is that the susceptibility loci for each disease are incomplete and the extent of the incompleteness varies between diseases. This prevents us, for example, from drawing any conclusions that 'AITD is less like AS than UC'. However, for the majority of the immune-mediated diseases listed in Table 2, we can assume that the common loci with the largest effects are known and that significant results here are true representations of AITD-disease overlap.

Downloaded from http://hmg.oxfordjournals.org/ at Istituto Dermopatico Immacolata on November 6, 2012

Table 1. Single locus test evidence for the seven newly identified AITD non-HLA susceptibility loci and 6q27 on the ImmunoChip.

Karyotype band	Most disease-associated SNP	Located within the gene	MAF in controls	Immune-mediated diseases previously associated with regions ($P < 5 \times 10^{-8}$) (reported SNP; r^2 with reported SNP in controls)	2285 GD patients and 9364 controls OR (95% CI)	*P*-value	462 HT patients and 9364 controls OR (95% CI)	*P*-value	Heterogeneity in disease association *P*-value	2282 GD, 451 HT patients and 9364 controls[a] OR (95% CI)	*P*-value
1p36.32	rs2843403 C>T	*MMEL1* (32 kb upstream of *TNFRSF1*)	0.362	CeD (26) (rs3748816/*TNFRSF14-MMEL1*; $r^2 = 0.99$) PBC (27) (rs3748816; $r^2 = 0.99$) RA (28) (rs3890745/*TNFRSF14-MMEL1*; $r^2 = 0.90$)	0.84 (0.79–0.90)	**7.94×10^{-7}**	0.97 (0.85–1.12)	0.696	0.0593	0.86 (0.81–0.92)	5.63×10^{-6}
2p25.1	rs1534422 A>G	216 kb upstream of *TRIB2*	0.455	Unconfirmed T1D locus[b]	1.16 (1.09–1.24)	4.69×10^{-6}	1.24 (1.08–1.41)	1.64×10^{-3}	0.394	1.17 (1.11–1.25)	**1.76×10^{-7}**
3q27.3/ 3q28	rs13093110 C>T	*LPP*	0.452	CeD (26) (rs1464510; $r^2 = 0.98$)	1.18 (1.10–1.26)	8.17×10^{-7}	1.20 (1.05–1.37)	7.09×10^{-3}	0.797	1.19 (1.12–1.26)	**3.69×10^{-8}**
6q15	rs72928038 G>A	*BACH2*	0.177	CD (29) (rs1847472; $r^2 = 0.39$) CeD (26) (rs10806425/*BACH2-MAP3K7*; $r^2 = 0.30$) T1D (21) (rs11755527; $r^2 = 0.19$)[c]	1.21 (1.12–1.32)	3.63×10^{-6}	1.30 (1.11–1.53)	1.36×10^{-3}	0.417	1.23 (1.14–1.32)	**1.23×10^{-7}**
6q27	imm_6_167338101 A>C	*FGFR1OP* (118 kb upstream of *CCR6*)	0.408	AITD (5) (rs9355610; $r^2 = 0.72$) CD (30) (rs2301436/*CCR6*; $r^2 = 0.61$) RA (31) (rs3093023/*CCR6*; $r^2 = 0.25$ Vitiligo (32) (rs2236313; $r^2 = 0.22$)	0.84 (0.79–0.90)	3.30×10^{-7}	0.88 (0.76–1.00)	0.0564	0.575	0.85 (0.80–0.90)	**1.64×10^{-7}**
11q21	rs4409785 T>C		0.173	Unconfirmed MS locus[d]	1.21 (1.11–1.31)	5.37×10^{-6}	1.34 (1.14–1.57)	3.54×10^{-4}	0.234	1.23 (1.14–1.33)	**7.69×10^{-8}**
12q12	rs4768412 C>T	*PRICKLE1*	0.363	CD (30) (rs11175593/*LRRK2-MUC19*)[d]	1.19 (1.11–1.27)	**3.30×10^{-7}**	1.00 (0.88–1.15)	0.949	0.0234[e]	1.16 (1.09–1.23)	3.56×10^{-6}
16p11.2	rs57348955 G>A	83 kb upstream of *ITGAM*	0.396	SLE (33) (rs1143679/*ITGAM*; $r^2 = 0.056$) T1D (13) (rs4988084/*IL27*)[f]	0.83 (0.77–0.89)	**3.76×10^{-8}**	0.91 (0.80–1.05)	0.188	0.190	0.84 (0.79–0.89)	5.13×10^{-8}

The reported SNPs have a *P*-value of $<1.12 \times 10^{-6}$ (see Materials and Methods) and were the most disease-associated SNP in the region with good genotype signal intensity plots (www.ImmunoBase.org). For each SNP, the most significant *P*-value is shown in bold. Note that the same controls were analysed in each analysis. AITD, autoimmune thyroid disease; MAF, minor allele frequency; HT, Hashimoto's thyroiditis; GD, Graves' disease; OR, Odds ratio for the minor allele; CI, confidence interval; CD, Crohn's disease; CeD, celiac disease; PBC, primary biliary cirrhosis; and SLE, systematic lupus erythematosus.

[a]We excluded 14 second degree relatives or closer between disease groups (see Materials and Methods).

[b]rs1534422 at 2p25.1 is an unconfirmed T1D risk locus (T1D $P = 2.1 \times 10^{-6}$; (13)).

[c]We used CEU data from the 1000 Genomes Project to estimate r^2 as the T1D SNP, rs11755527, was not available in the ImmunoChip data.

[d]rs4409785 at 11q21 is an unconfirmed MS risk locus [MS $P = 6.3 \times 10^{-7}$; (14)].

[e]Assuming 13 [8 (Table 1) and 5 (Supplementary Material, Table S1)] independent tests, the adjusted *P*-value was 3.85×10^{-3} for the 0.05 level of significance based on the Bonferroni correction for multiple testing.

[f]No available genotype data which included both SNPs.

Downloaded from http://hmg.oxfordjournals.org/ at Istituto Dermopatico Immacolata on November 6, 2012

Table 2. The overlap between AITD and the 11 other immune-mediated diseases that provided susceptibility loci to the ImmunoChip. A negative *Z*-value indicates that SNPs within disease-associated regions were more likely to be associated with AITD than SNPs within regions not known to be associated with the disease. A positive *Z*-value indicates that SNPs within regions not known to be associated with the disease were more likely to be associated with AITD than SNPs within associated regions.

Immune-mediated disease	Number of non-HLA loci	Number of SNPs within loci	Wilcoxon rank test *Z*	*P*
AS	8	2079	3.593	3.27×10^{-4}
CeD	30	9866	0.147	0.883
CD	66	22 585	2.242	0.0250
IgA deficiency	1	122	−1.208	0.227
MS	26	7147	−0.277	0.782
PBC	1	209	−0.840	0.401
PS	25	6767	2.872	4.08×10^{-3}
RA	25	7742	−3.992	6.56×10^{-5}
SLE	35	9099	−0.726	0.468
T1D	41	12 658	−2.146	0.0327
UC	28	8901	3.140	1.69×10^{-3}

DISCUSSION

We have found evidence for seven risk loci not previously associated with AITD and expect that even more AITD risk loci will be discovered as more AITD patients are genotyped on the ImmunoChip. In this study, as the ImmunoChip contains densely mapped immune-mediated disease loci representing only a small proportion, ∼1%, of the genome, we used a permutation approach to derive a significance level rather than adopting a genome-wide significance level. Importantly, regardless of which discovery significance level was adopted, confirmation of the new AITD risk loci will require independent replication evidence. We found some statistical evidence that AITD, outside of the HLA, has a shared genetic architecture with both RA and T1D. AITD often co-exists with other immune-mediated diseases and a recent study estimating the prevalence of these co-existing disorders found that RA was the most common, occurring in 3.15% of GD patients and 4.24% of HT patients studied; T1D occurred in 1.11% of GD patients and 1.01% of HT patients (17). We note that the prevalence study (17) included vast majority of our AITD patients, and that AITD patients were recruited from endocrine clinics and not rheumatology clinics.

The implicated candidate genes for the established and seven new AITD risk loci indicate the importance of T lymphocyte signalling (*PTPN22*), T-regulatory cell function (*CTLA4* and *IL2RA*), lymphocyte trafficking (*CCR6*) and a newly identified possible link with the anti-viral immune response (*BACH2*). Nevertheless, other AITD-associated regions suggest candidate genes of uncertain or unknown functions or roles in autoimmunity of AITD, such as *LPP* or *TRIB2*. Two regions, 11q21 and 12q12, have no annotated protein-coding candidate genes, indicating how much of the aetiology of this and other immune-mediated diseases remain to be explained.

MATERIALS AND METHODS

Subjects

The British AITD cases consisted of 2374 (390 males and 1984 females) GD and 474 (69 males and 405 females) HT patients (18). The British control population consisted of 6894 subjects drawn from the British 1958 Birth Cohort (1958BC; http://wwww.cls.ioe.ac.uk/studies.asp?section=000100020003) and 3059 subjects drawn from the UK Blood Services Common Control Collection (UKBS-CC) (6,19). All subjects were of white European ancestry with written informed consent and Ethics Committee/Institutional Review Board approval.

SNP selection

The dense SNP map for 186 confirmed and distinct loci, mainly derived from GWA studies, which have been associated (required $P < 5 \times 10^{-8}$) with one or more autoimmune diseases, consisted of all known SNPs, and small insertion deletions, in the dbSNP database, in the 1000 Genomes project (February 2010 release) and in additional sequencing data provided by collaborators. The ImmunoChip contains 196 524 polymorphisms (718 insertion/deletions and 195 806 SNPs) (3,20).

Genotyping

Samples were genotyped using a custom Illumina 200K Infinium high-density array according to the manufacturer's protocol at the Wellcome Trust Sanger Institute (WTSI) in Hinxton, UK, and University of Virginia (UVA) in Charlottesville, USA. In WTSI, 2374 GD, 474 HT, 1478 1958BC and 3059 UKBS-CC samples were processed, and in UVA, 5416 1958BC samples were processed.

Sample quality control

Samples were excluded based on per-sample call rate (Supplementary Material, Figs S9–S11), outlying autosomal heterozygosity (Supplementary Material, Figs S9–S11), inconsistent recorded and genotype-inferred sex, non-European ancestry (21) (Supplementary Material, Figs S12–S14), duplication and being closely related to another sample in the study. Relatedness between the study participants was estimated by a identity by state (IBS) statistic, we defined duplicates, or monozygotic twins, as having an IBS >0.98 and second degree relatives or closer, as having an IBS >0.1875 (22); the sample with the lower call rate was excluded from further analyses. Finally, we checked control sample's identity using the 169 intentional duplicates between the WTSI and the UVA, and using in-house genotype data. We identified 256 samples, which had incorrect identifiers as a result of manifest misreading, plate rotation and chip-related issues, and excluded 95 samples from unknown subjects (i.e. could not be assigned to identifiers). In addition, we found ten additional duplicates between the genotyping centers. All potential sample identity errors were discussed with the relevant genotyping centers by Neil M. Walker. After sample quality control, 2285 GD, 462 HT, 6541 1958BC and

Downloaded from http://hmg.oxfordjournals.org/ at Istituto Dermopatico Immacolata on November 6, 2012

2823 UKBS-CC samples were available for analysis. When we analysed GD and HT patients together, we excluded 14 second degree relatives or closer between the disease groups, which resulted in 2282 GD and 451 HT samples.

SNP quality control

SNP genotypes were called using Illumina GenomeStudio GenTrain 2.0 algorithm and genotypes with a gencall score (a quality metric for the distance a sample is from the center of the nearest genotype cluster) <0.15 were not assigned genotypes. SNPs were excluded if the MAF fell $<5\%$ in controls, if they deviated from Hardy–Weinberg equilibrium (HWE; z-score >5) in controls or if the per-SNP call rate fell $<95\%$ in cases or controls. We note that SNP quality control was performed in WTSI and UVA-processed controls separately. In addition, for SNPs from the pseudoautosomal regions on chromosome X, we tested for deviation from HWE in male and female controls separately after discovering SNP genotype frequency differences between sexes in patients and controls (*e.g.* rs306891). We excluded SNPs from the HLA (NCBI build 36 coordinates chr6:29690000-33498585).

Previously, despite differences in subject's age and in DNA processing between 1958BC and UKBS-CC control subjects, few differences in SNP allele frequencies between 1958BC and UKBS-CC controls were observed (19), which justified the combination of these two control groups into a single control group. To assess the possible differential genotyping errors between the WTSI and UVA genotyping centers, we tested for differences in SNP allele frequencies in controls between genotyping centers; the test was stratified by geographical region of Great Britain (23). The quantile–quantile plot for the WTSI control versus UVA control 1 degree-of-freedom (1-df) association tests (Supplementary Material, Fig. S15) showed a slight inflation of the test statistic [inflation factor (λ) was 1.017; as the inflation factor scales with sample size, a more informative and comparable measure of inflation is the inflation factor for an equivalent study of 1000 cases and 1000 controls, λ_{1000} (24), which was 1.004 for the WTSI control/UVA control analysis]. As the ImmunoChip was designed to include all known and genotype-able SNPs across 186 distinct immune-mediated disease susceptibility loci, there was inflation in the test statistic for the AITD patient versus control association tests: HT patient/control $\lambda_{1000} = 1.06$ ($\lambda = 1.05$); GD patient/control $\lambda_{1000} = 1.08$ ($\lambda = 1.28$); and HT and GD patient/control $\lambda_{1000} = 1.08$ ($\lambda = 1.33$). We note that λ was based on non-HLA common SNPs from regions not previously associated with AITD (excluded regions reported in Supplementary Material, Table S1) that passed quality control. We also estimated the inflation factor based on the WTCCC replication SNPs for reading and maths ability, and psychosis endophenotypes: HT patient/control $\lambda_{1000} = 1.02$ ($\lambda = 1.02$); GD patient/control $\lambda_{1000} = 1.03$ ($\lambda = 1.11$); and HT and GD patient/control $\lambda_{1000} = 1.04$ ($\lambda = 1.17$). We note that 2605 WTCCC replication SNPs passed quality control and had a MAF >0.05, and that the observed inflation of the test statistic was a result of the replication SNPs overlapping some of the immune-mediated disease loci, for example, the T1D loci *COBL*/7p12.1 and *GLIS3*/9p24.2. As the inflation factor λ_{1000} of the test statistics from the WTSI control/UVA control analysis was just over 1 and we found a similar inflation of the test statistics for GD patient/controls and for HT patient/controls at 6–8% in the main analysis and at 2–3% in the subset of WTCCC replication SNPs, population structure is unlikely to a major source of inflation. After SNP quality control, 103 875 common SNPs were available for analysis.

Statistical analysis

All statistical analyses were performed in R (http://www.r-project.org) using the snpStats package available from the Bioconductor project (http://www.bioconductor.org). After applying sample and SNP quality control exclusions, we analysed the case/control data using the 1-df Cochran-Armitage trend test. SNPs were modelled using a numerical indicator variable coded 0, 1 and 2, representing the occurrences of the minor allele and assuming multiplicative allelic effects. We also tested the disease-associated SNPs reported in Table 1 for differences between the multiplicative allelic effects model and the full genotype model, and no significant differences were found. We analysed SNPs on the X-chromosome using the method proposed by Clayton (25).

We initially analysed the GD and HT patients separately, but against the same controls. Before combining the GD and HT patients into a single patient group, as we had analysed GD and HT patients against the same controls, we tested for disease association heterogeneity between the two AITDs by testing for differences in SNP allele frequencies between GD and HT patient groups. In the absence of disease association heterogeneity, we proceeded to combine the GD and HT patients and re-analysed.

As GD and HT patients, unlike the controls, were not recruited from all 12 geographical regions of Great Britain (seven and six regions, respectively), we did not stratify the analysis by geographical region to allow for population structure. We note that previously, little difference was reported between analyses with or without regional stratification (19). However, to check that we were gaining power with the additional controls rather than false-positive findings through ignoring the population structure, we tested for population heterogeneity in SNP genotype frequencies across the regions to identify SNPs that differentiate between the regions (Supplementary Material, Table S2). None of the seven new or previously reported AITD risk loci showed large allele frequency differences between controls from the 12 regions.

To adjust P-values for the multiple tests and to allow for the correlation between SNP genotypes, we adopted a permutation test approach. We performed 100 000 permutation tests of the common SNPs on the ImmunoChip, permuting the disease status label for each test, to provide the distribution of the test statistic under the null hypothesis of no association between SNP genotypes and disease status. The adjusted P-value for 0.05 was 1.12×10^{-6}.

To examine the overlap between AITD and the other 11 immune-mediated diseases that provided susceptibility loci to the ImmunoChip, we tested whether the distribution of AITD test statistics varied between SNPs within regions associated with another immune-mediated disease and within

Downloaded from http://hmg.oxfordjournals.org/ at Istituto Dermopatico Immacolata on November 6, 2012

regions not associated with that disease. We adopted a gene set enrichment approach based on the Wilcoxon rank test as implemented in the R package wgsea (Chris Wallace, www-gene.cimr.cam.ac.uk/staff/wallace) (16). SNPs within the non-HLA immune-mediated disease loci were divided into two groups for each immune-mediated disease: (i) SNPs within regions associated with disease at the time the ImmunoChip was designed; and (ii) SNPs within regions not associated with that disease. We then tested the null hypothesis that the distribution of *P*-values describing the evidence for association of each SNP with AITD was the same in the two groups using a Wilcoxon rank test with a permutation-based variance estimate to allow for the considerable between SNP correlation (16). We used 10 000 permutations to estimate the variance of the Wilcoxon rank test.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We would like to thank all AITD patients and control subjects for participating in this study. We thank nurses and doctors for recruiting AITD patients into the AITD National Collection. We thank the members of each disease consortium who initiated and sustained the cross-disease ImmunoChip project. We thank Jeffrey Barrett for assistance with ImmunoChip SNP selection and for ImmunoChip-related correspondence. We thank Jennifer Stone for co-ordinating the ImmunoChip design and production at Illumina. The authors wish to acknowledge Sarah Edkins, Emma Gray, Doug Simpkin, Sarah Hunt and staff of the Sanger Institute's Sample Logistics, Genotyping and Variation Informatics facilities, respectively. The Diabetes and Inflammation Laboratory is funded by the Juvenile Diabetes Research Foundation International, the Wellcome Trust and the National Institute for Health Research Cambridge Biomedical Centre. The Cambridge Institute for Medical Research (CIMR) is in receipt of a Wellcome Trust Strategic Award (079895). Core infrastructure support to the Wellcome Trust Centre for Human Genetics, Oxford was provided by grant 090532/Z/09/Z. The WTCCC is funded by Wellcome Trust awards 076113 and 083948. CW is supported by the Wellcome Trust (089989). We acknowledge use of DNA from The UK Blood Services collection of Common Controls (UKBS-CC collection), which is funded by the Wellcome Trust grant 076113/C/04/Z and by the USA National Institute for Health Research program grant to the National Health Service Blood and Transplant (RP-PG-0310-1002). We acknowledge the use of DNA from the British 1958 Birth Cohort collection, which is funded by the UK Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. This research utilized 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, the National Institute of Allergy and Infectious Diseases, the National Human Genome Research Institute, the National Institute of Child Health and Human Development and the Juvenile Diabetes Research Foundation International and is supported by the USA National Institutes of Health grant U01-DK062418.

Conflict of Interest statement. None declared.

FUNDING

The research leading to these results has received funding from the European Union's 7th Framework Programme (FP7/2007–2013) under grant agreement No. 241447 [Natural immunomodulators as novel immunotherapies for type 1 diabetes (NAIMIT)]. Funding to pay the Open Access publication charges for this article was provided by the Wellcome Trust.

REFERENCES

1. Smyth, D., Cooper, J.D., Collins, J.E., Heward, J.M., Franklyn, J.A., Howson, J.M., Vella, A., Nutland, S., Rance, H.E., Maier, L. *et al.* (2004) Replication of an association between the lymphoid tyrosine phosphatase locus (LYP/PTPN22) with type 1 diabetes, and evidence for its role as a general autoimmunity locus. *Diabetes*, **53**, 3020–3023.
2. Ueda, H., Howson, J.M., Esposito, L., Heward, J., Snook, H., Chamberlain, G., Rainbow, D.B., Hunter, K.M., Smith, A.N., Di Genova, G. *et al.* (2003) Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. *Nature*, **423**, 506–511.
3. Trynka, G., Hunt, K.A., Bockett, N.A., Romanos, J., Mistry, V., Szperl, A., Bakker, S.F., Bardella, M.T., Bhaw-Rosun, L., Castillejo, G. *et al.* (2011) Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat. Genet.*, **43**, 1193–1201.
4. McGrogan, A., Seaman, H.E., Wright, J.W. and de Vries, C.S. (2008) The incidence of autoimmune thyroid disease: a systematic review of the literature. *Clin. Endocrinol. (Oxf.)*, **69**, 687–696.
5. Chu, X., Pan, C.M., Zhao, S.X., Liang, J., Gao, G.Q., Zhang, X.M., Yuan, G.Y., Li, C.G., Xue, L.Q., Shen, M. *et al.* (2011) A genome-wide association study identifies two new risk loci for Graves' disease. *Nat. Genet.*, **43**, 897–901.
6. Burton, P.R., Clayton, D.G., Cardon, L.R., Craddock, N., Deloukas, P., Duncanson, A., Kwiatkowski, D.P., McCarthy, M.I., Ouwehand, W.H., Samani, N.J. *et al.* (2007) Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. *Nat. Genet.*, **39**, 1329–1337.
7. Kochi, Y., Yamada, R., Suzuki, A., Harley, J.B., Shirasawa, S., Sawada, T., Bae, S.C., Tokuhiro, S., Chang, X., Sekine, A. *et al.* (2005) A functional variant in FCRL3, encoding Fc receptor-like 3, is associated with rheumatoid arthritis and several autoimmunities. *Nat. Genet.*, **37**, 478–485.
8. Simmonds, M.J., Yesmin, K., Newby, P.R., Brand, O.J., Franklyn, J.A. and Gough, S.C. (2010) Confirmation of association of chromosome 5q31–33 with United Kingdom Caucasian Graves' disease. *Thyroid*, **20**, 413–417.
9. Song, H.D., Liang, J., Shi, J.Y., Zhao, S.X., Liu, Z., Zhao, J.J., Peng, Y.D., Gao, G.Q., Tao, J., Pan, C.M. *et al.* (2009) Functional SNPs in the SCGB3A2 promoter are associated with susceptibility to Graves' disease. *Hum. Mol. Genet.*, **18**, 1156–1170.
10. Brand, O.J., Lowe, C.E., Heward, J.M., Franklyn, J.A., Cooper, J.D., Todd, J.A. and Gough, S.C. (2007) 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.)*, **66**, 508–512.
11. Chistiakov, D.A., Chistiakova, E.I., Voronova, N.V., Turakulov, R.I. and Savost'anov, K.V. (2011) A variant of the Il2ra/Cd25 gene predisposing to graves' disease is associated with increased levels of soluble interleukin-2 receptor. *Scand. J. Immunol.*, **74**, 496–501.
12. Hong, S.W., Kim, S. and Lee, D.K. (2008) The role of Bach2 in nucleic acid-triggered antiviral innate immune responses. *Biochem. Biophys. Res. Commun.*, **365**, 426–432.
13. Barrett, J.C., Clayton, D.G., Concannon, P., Akolkar, B., Cooper, J.D., Erlich, H.A., Julier, C., Morahan, G., Nerup, J., Nierras, C. *et al.* (2009)

Downloaded from http://hmg.oxfordjournals.org/ at Istituto Dermopatico Immacolata on November 6, 2012

Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat. Genet.*, **41**, 703–707.
14. Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C.C., Patsopoulos, N.A., Moutsianas, L., Dilthey, A., Su, Z., Freeman, C., Hunt, S.E. *et al.* (2011) Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature*, **476**, 214–219.
15. Brand, O.J. and Gough, S.C. (2010) Genetics of thyroid autoimmunity and the role of the TSHR. *Mol. Cell Endocrinol.*, **322**, 135–143.
16. Heinig, M., Petretto, E., Wallace, C., Bottolo, L., Rotival, M., Lu, H., Li, Y., Sarwar, R., Langley, S.R., Bauerfeind, A. *et al.* (2010) A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk. *Nature*, **467**, 460–464.
17. Boelaert, K., Newby, P.R., Simmonds, M.J., Holder, R.L., Carr-Smith, J.D., Heward, J.M., Manji, N., Allahabadia, A., Armitage, M., Chatterjee, K.V. *et al.* (2010) Prevalence and relative risk of other autoimmune diseases in subjects with autoimmune thyroid disease. *Am. J. Med.*, **123**, 183, e181–189.
18. Manji, N., Carr-Smith, J.D., Boelaert, K., Allahabadia, A., Armitage, M., Chatterjee, V.K., Lazarus, J.H., Pearce, S.H., Vaidya, B., Gough, S.C. *et al.* (2006) Influences of age, gender, smoking, and family history on autoimmune thyroid disease phenotype. *J. Clin. Endocrinol. Metab.*, **91**, 4873–4880.
19. 2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, **447**, 661–678.
20. Cortes, A. and Brown, M.A. (2011) Promise and pitfalls of the Immunochip. *Arthritis Res. Ther.*, **13**, 101.
21. Cooper, J.D., Smyth, D.J., Smiles, A.M., Plagnol, V., Walker, N.M., Allen, J.E., Downes, K., Barrett, J.C., Healy, B.C., Mychaleckyj, J.C. *et al.* (2008) Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. *Nat. Genet.*, **40**, 1399–1401.
22. Anderson, C.A., Pettersson, F.H., Clarke, G.M., Cardon, L.R., Morris, A.P. and Zondervan, K.T. (2010) Data quality control in genetic case-control association studies. *Nat. Protoc.*, **5**, 1564–1573.
23. Clayton, D.G., Walker, N.M., Smyth, D.J., Pask, R., Cooper, J.D., Maier, L.M., Smink, L.J., Lam, A.C., Ovington, N.R., Stevens, H.E. *et al.* (2005) Population structure, differential bias and genomic control in a large-scale, case–control association study. *Nat. Genet.*, **37**, 1243–1246.
24. Freedman, M.L., Reich, D., Penney, K.L., McDonald, G.J., Mignault, A.A., Patterson, N., Gabriel, S.B., Topol, E.J., Smoller, J.W., Pato, C.N. *et al.* (2004) Assessing the impact of population stratification on genetic association studies. *Nat. Genet.*, **36**, 388–393.
25. Clayton, D.G. (2009) Sex chromosomes and genetic association studies. *Genome Med.*, **1**, 110.
26. Dubois, P.C., Trynka, G., Franke, L., Hunt, K.A., Romanos, J., Curtotti, A., Zhernakova, A., Heap, G.A., Adany, R., Aromaa, A. *et al.* (2010) Multiple common variants for celiac disease influencing immune gene expression. *Nat. Genet.*, **42**, 295–302.
27. Hirschfield, G.M., Liu, X., Han, Y., Gorlov, I.P., Lu, Y., Xu, C., Chen, W., Juran, B.D., Coltescu, C., Mason, A.L. *et al.* (2010) Variants at IRF5-TNPO3, 17q12–21 and MMEL1 are associated with primary biliary cirrhosis. *Nat. Genet.*, **42**, 655–657.
28. Kurreeman, F.A., Stahl, E.A., Okada, Y., Liao, K., Diogo, D., Raychaudhuri, S., Freudenberg, J., Kochi, Y., Patsopoulos, N.A., Gupta, N. *et al.* (2012) Use of a multiethnic approach to identify rheumatoid-arthritis-susceptibility loci, 1p36 and 17q12. *Am. J. Hum. Genet.*, **90**, 524–532.
29. Franke, A., McGovern, D.P., Barrett, J.C., Wang, K., Radford-Smith, G.L., Ahmad, T., Lees, C.W., Balschun, T., Lee, J., Roberts, R. *et al.* (2010) Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat. Genet.*, **42**, 1118–1125.
30. Barrett, J.C., Hansoul, S., Nicolae, D.L., Cho, J.H., Duerr, R.H., Rioux, J.D., Brant, S.R., Silverberg, M.S., Taylor, K.D., Barmada, M.M. *et al.* (2008) Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.*, **40**, 955–962.
31. Stahl, E.A., Raychaudhuri, S., Remmers, E.F., Xie, G., Eyre, S., Thomson, B.P., Li, Y., Kurreeman, F.A., Zhernakova, A., Hinks, A. *et al.* (2010) Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat. Genet.*, **42**, 508–514.
32. Quan, C., Ren, Y.Q., Xiang, L.H., Sun, L.D., Xu, A.E., Gao, X.H., Chen, H.D., Pu, X.M., Wu, R.N., Liang, C.Z. *et al.* (2010) Genome-wide association study for vitiligo identifies susceptibility loci at 6q27 and the MHC. *Nat. Genet.*, **42**, 614–618.
33. Nath, S.K., Han, S., Kim-Howard, X., Kelly, J.A., Viswanathan, P., Gilkeson, G.S., Chen, W., Zhu, C., McEver, R.P., Kimberly, R.P. *et al.* (2008) A nonsynonymous functional variant in integrin-alpha(M) (encoded by ITGAM) is associated with systemic lupus erythematosus. *Nat. Genet.*, **40**, 152–154.

Downloaded from http://hmg.oxfordjournals.org/ at Istituto Dermopatico Immacolata on November 6, 2012

Supplementary text, tables and figures for "Seven newly identified loci for autoimmune thyroid disease"

Jason D Cooper, Matthew J Simmonds, Neil M Walker, Oliver Burren, Oliver J Brand, Hui Guo, Chris Wallace, Helen Stevens, Gillian Coleman, Wellcome Trust Case Control Consortium, Jayne A Franklyn, John A Todd and Stephen CL Gough

Members of the WTCCC involved in this study

Jan Aerts[1], Tariq Ahmad[2], Hazel Arbury[1], Anthony Attwood[1,3,4], Adam Auton[5], Stephen G Ball[6], Anthony J Balmforth[6], Chris Barnes[1], Jeffrey C Barrett[1], Inês Barroso[1], Anne Barton[7], Amanda J Bennett[8], Sanjeev Bhaskar[1], Katarzyna Blaszczyk[9], John Bowes[7], Oliver J Brand[8,10], Peter S Braund[11], Francesca Bredin[12], Gerome Breen[13,14], Morris J Brown[15], Ian N Bruce[7], Jaswinder Bull[16], Oliver S Burren[17], John Burton[1], Jake Byrnes[18], Sian Caesar[19], Niall Cardin[5], Chris M Clee[1], Alison J Coffey[1], John MC Connell[20], Donald F Conrad[1], Jason D Cooper[17], Anna F Dominiczak[20], Kate Downes[17], Hazel E Drummond[21], Darshna Dudakia[16], Andrew Dunham[1], Bernadette Ebbs[16], Diana Eccles[22], Sarah Edkins[1], Cathryn Edwards[23], Anna Elliot[16], Paul Emery[24], David M Evans[25], Gareth Evans[26], Steve Eyre[7], Anne Farmer[14], I Nicol Ferrier[27], Edward Flynn[7], Alistair Forbes[28], Liz Forty[29], Jayne A Franklyn[10,30], Timothy M Frayling[2], Rachel M Freathy[2], Eleni Giannoulatou[5], Polly Gibbs[16], Paul Gilbert[7], Katherine Gordon-Smith[19,29], Emma Gray[1], Elaine Green[29], Chris J Groves[8], Detelina Grozeva[29], Rhian Gwilliam[1], Anita Hall[16], Naomi Hammond[1], Matt Hardy[17], Pile Harrison[31], Neelam Hassanali[8], Husam Hebaishi[1], Sarah Hines[16], Anne Hinks[7], Graham A Hitman[32], Lynne Hocking[33], Chris Holmes[5], Eleanor Howard[1], Philip Howard[34], Joanna MM Howson[17], Debbie Hughes[16], Sarah Hunt[1], John D Isaacs[35], Mahim Jain[18], Derek P Jewell[36], Toby Johnson[34], Jennifer D Jolley[3,4], Ian R Jones[29], Lisa A Jones[19], George Kirov[29], Cordelia F Langford[1], Hana Lango-Allen[2], G Mark Lathrop[37], James Lee[12], Kate L Lee[34], Charlie Lees[21], Kevin Lewis[1], Cecilia M Lindgren[8,18], Meeta Maisuria-Armer[17], Julian Maller[18], John Mansfield[38], Jonathan L Marchini[5], Paul Martin[7], Dunecan CO Massey[12], Wendy L McArdle[39], Peter McGuffin[14], Kirsten E McLay[1], Gil McVean[5,18], Alex Mentzer[40], Michael L Mimmack[1], Ann E Morgan[41], Andrew P Morris[18], Craig Mowat[42], Patricia B Munroe[34], Simon Myers[18], William Newman[26], Elaine R Nimmo[21], Michael C O'Donovan[29], Abiodun Onipinla[34], Nigel R Ovington[17], Michael J Owen[29], Kimmo Palin[1], Aarno Palotie[1], Kirstie Parnel'[2], Richard Pearson[8], David Pernet[16], John RB Perry[2,18], Anne Phillips[42], Vincent Plagnol[17], Natalie J Prescott[9], Inga Prokopenko[8,18], Michael A Quail[1], Suzanne Rafelt[11], Nigel W Rayner[8,18], David M Reid[33], Anthony Renwick[16], Susan M Ring[39], Neil Robertson[8,18], Samuel Robson[1], Ellie Russell[29], David St Clair[13], Jennifer G Sambrook[3,4], Jeremy D Sanderson[40], Stephen J Sawcer[43], Helen Schuilenburg[17], Carol E Scott[1], Richard Scott[16], Sheila Seal[16], Sue Shaw-Hawkins[34], Beverley M Shields[2], Matthew J Simmonds[8,10], Debbie J Smyth[17], Elilan Somaskantharajah[1], Katarina Spanova1[6], Sophia Steer[44], Jonathan Stephens[3,4], Helen E Stevens[17], Kathy Stirrups[1], Millicent A Stone[45,46], David P Strachan[47], Zhan Su[5], Deborah PM Symmons[7], John R Thompson[48], Wendy Thomson[7], Martin D Tobin[48], Mary E Travers[8], Clare Turnbull[16], Damjan Vukcevic[18], Louise V Wain[48], Mark Walker[49], Neil M Walker[17], Chris Wallace[17], Margaret Warren-Perry[16], Nicholas A Watkins[3,4], John Webster[50], Michael N Weedon[2], Anthony G Wilson[51], Matthew Woodburn[17], B Paul Wordsworth[52], Chris Yau[5], Allan H Young[27,53], Eleftheria Zeggini[1], Matthew A Brown[52,54], Paul R Burton[48], Mark J Caulfield[34], Alastair Compston[43], Martin Farrall[55], Stephen CL Gough[8,10,30], Alistair S Hall[6], Andrew T Hattersley[2,56], Adrian VS Hill[18], Christopher G Mathew[9], Marcus Pembrey[57], Jack Satsangi[21], Michael R Stratton[1,16], Jane Worthington[7], Matthew E Hurles[1], Audrey Duncanson[58], Willem H Ouwehand[1,3,4], Miles Parkes[12], Nazneen Rahman[16], John A Todd[17], Nilesh J Samani1[1,59], Dominic P Kwiatkowski[1,18], Mark I McCarthy[8,18,60], Nick Craddock[29], Panos Deloukas[1], Peter Donnelly[5,18].

[1] The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA UK.

[2] Genetics of Complex Traits, Peninsula College of Medicine and Dentistry University of Exeter, EX1 2LU, UK.

[3] Department of Haematology, University of Cambridge, Long Road, Cambridge, CB2 0PT, UK.

[4] National Health Service Blood and Transplant, Cambridge Centre, Long Road, Cambridge CB2 0PT, UK.

[5] Department of Statistics, University of Oxford, 1 South Parks Road, Oxford, OX1 3TG, UK.

[6] Multidisciplinary Cardiovascular Research Centre (MCRC), Leeds Institute of Genetics, Health and Therapeutics (LIGHT), University of Leeds, Leeds, LS2 9JT, UK.

[7] arc Epidemiology Unit, Stopford Building, University of Manchester, Oxford Road, Manchester, M13 9PT, UK.

[8] Oxford Centre for Diabetes, Endocrinology and Medicine, University of Oxford, Churchill Hospital, Oxford OX3 7LJ, UK.

[9] Department of Medical and Molecular Genetics, King's College London School of Medicine, 8th Floor Guy's Tower, Guy's Hospital, London, SE1 9RT, UK.

[10] Centre for Endocrinology, Diabetes and Metabolism, Institute of Biomedical Research, University of Birmingham, Birmingham, B15 2TT, UK.

[11] Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Groby Road, Leicester LE3 9QP, UK.

[12] IBD Genetics Research Group, Addenbrooke's Hospital, Cambridge, CB2 0QQ, UK.

[13] University of Aberdeen, Institute of Medical Sciences, Foresterhill, Aberdeen AB25 2ZD, UK.

[14] SGDP, The Institute of Psychiatry, King's College London, De Crespigny Park, Denmark Hill, London SE5 8AF, UK.

[15] Clinical Pharmacology Unit, University of Cambridge, Addenbrookes Hospital, Hills Road, Cambridge CB2 2QQ, UK.

[16] Section of Cancer Genetics, Institute of Cancer Research, 15 Cotswold Road, Sutton SM2 5NG, UK.

[17] Juvenile Diabetes Research Foundation/Wellcome Trust Diabetes and Inflammation

Laboratory, Department of Medical Genetics, Cambridge Institute for Medical Research, University of Cambridge, Wellcome Trust/MRC Building, Cambridge CB2 0XY, UK.

[18] The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK.

[19] Department of Psychiatry, University of Birmingham, National Centre for Mental Health, 25 Vincent Drive, Birmingham, B15 2FG, UK.

[20] BHF Glasgow Cardiovascular Research Centre, University of Glasgow, 126 University Place, Glasgow, G12 8TA, UK.

[21] Gastrointestinal Unit, Division of Medical Sciences, School of Molecular and Clinical Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK.

[22] Academic Unit of Genetic Medicine, University of Southampton, Southampton, UK.

[23] Endoscopy Regional Training Unit, Torbay Hospital, Torbay TQ2 7AA, UK.

[24] Academic Unit of Musculoskeletal Disease, University of Leeds, Chapel Allerton Hospital, Leeds, West Yorkshire LS7 4SA, UK.

[25] MRC Centre for Causal Analyses in Translational Epidemiology, Department of Social Medicine, University of Bristol, Bristol, BS8 2BN, UK.

[26] Department of Medical Genetics, Manchester Academic Health Science Centre (MAHSC), University of Manchester, Manchester M13 0JH, UK.

[27] School of Neurology, Neurobiology and Psychiatry, Royal Victoria Infirmary, Queen Victoria Road, Newcastle upon Tyne, NE1 4LP, UK.

[28] Institute for Digestive Diseases, University College London Hospitals Trust, London NW1 2BU, UK.

[29] MRC Centre for Neuropsychiatric Genetics and Genomics, School of Medicine, Cardiff University, Heath Park, Cardiff, CF14 4XN, UK.

[30] University Hospital Birmingham NHS Foundation Trust, Birmingham, B15 2TT, UK.

[31] University of Oxford, Institute of Musculoskeletal Sciences, Botnar Research Centre, Oxford, OX3 7LD, UK.

[32] Centre for Diabetes and Metabolic Medicine, Barts and The London, Royal London Hospital, Whitechapel, London, E1 1BB, UK.

[33] Bone Research Group, Department of Medicine and Therapeutics, University of Aberdeen, Aberdeen, AB25 2ZD, UK.

[34] Clinical Pharmacology and Barts and The London Genome Centre, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK.

[35] Institute of Cellular Medicine, Musculoskeletal Research Group, 4th Floor, Catherine Cookson Building, The Medical School, Framlington Place, Newcastle upon Tyne, NE2 4HH, UK.

[36] Gastroenterology Unit, Radcliffe Infirmary, University of Oxford, Oxford, OX2 6HE, UK.

[37] Centre National de Genotypage, 2, Rue Gaston Cremieux, Evry, Paris 91057, France.

[38] Department of Gastroenterology & Hepatology, University of Newcastle upon Tyne, Royal Victoria Infirmary, Newcastle upon Tyne NE1 4LP, UK.

[39] ALSPAC Laboratory, Department of Social Medicine, University of Bristol, BS8 2BN, UK.

[40] Division of Nutritional Sciences, King's College London School of Biomedical and Health Sciences, London SE1 9NH, UK.

[41] NIHR-Leeds Musculoskeletal Biomedical Research Unit, University of Leeds, Chapel Allerton Hospital, Leeds, West Yorkshire LS7 4SA, UK.

[42] Department of General Internal Medicine, Ninewells Hospital and Medical School, Ninewells Avenue, Dundee DD1 9SY, UK.

[43] Department of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge, CB2 2QQ, UK.

[44] Clinical and Academic Rheumatology, Kings College Hosptal National Health Service Foundation Trust, Denmark Hill, London SE5 9RS, UK.

[45] University of Toronto, St. Michael's Hospital, 30 Bond Street, Toronto, Ontario M5B 1W8, Canada.

[46] University of Bath, Claverdon, Norwood House, Room 5.11a Bath Somerset BA2 7AY, UK.

[47] Division of Community Health Sciences, St George's, University of London, London SW17 0RE, UK.

[48] Departments of Health Sciences and Genetics, University of Leicester, 217 Adrian Building, University Road, Leicester, LE1 7RH, UK.

[49] Diabetes Research Group, School of Clinical Medical Sciences, Newcastle University, Framlington Place, Newcastle upon Tyne NE2 4HH, UK.

[50] Medicine and Therapeutics, Aberdeen Royal Infirmary, Foresterhill, Aberdeen, Grampian AB9 2ZB, UK.

[51] School of Medicine and Biomedical Sciences, University of Sheffield, Sheffield, S10 2JF, UK.

[52] Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, Nuffield Orthopaedic Centre, University of Oxford, Windmill Road, Headington, Oxford, OX3 7LD, UK.

[53] UBC Institute of Mental Health, 430-5950 University Boulevard Vancouver, British Columbia, V6T 1Z3, Canada.

Supplementary Tables

Table S1: Evidence for reported AITD non-HLA susceptibility loci. Genotype signal intensity cluster plots have been examined visually. Note that the AITD locus *SCGB3A2*/5q32 was not included on the ImmunoChip. Hashimoto's thyroiditis (HT), Graves' disease (GD), minor allele frequency (MAF), odds ratio (OR) and confidence interval (CI).

			2,285 GD patients and 9,364 controls		462 HT patients and 9,364 controls		Heterogeneity in disease association	2,282 GD, 451 HT and 9,364 controls	
Gene Karyotype band	Most associated SNP	MAF in in controls	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value	*P*-value	OR (95% CI)	*P*-value
PTPN22 1p13.2	rs2476601 G>A	0.0962	1.55 (1.41-1.71)	$4.03x10^{-16}$	2.02 (1.69-2.41)	$3.74x10^{-15}$	†$6.76x10^{-3}$	1.63 (1.49-1.78)	$9.69x10^{-23}$
FCRL3 1q23.1	rs7522061 T>C	0.480	1.16 (1.08-1.23)	$1.08x10^{-5}$	1.03 (0.90-1.18)	0.634	0.122	1.14 (1.07-1.21)	$2.36x10^{-5}$
CTLA4/ICOS 2q33.2	rs11571297 G>A	0.493	0.72 (0.67-0.77)	$2.81x10^{-23}$	0.82 (0.72-0.94)	$3.21x10^{-3}$	0.0682	0.73 (0.69-0.78)	$2.09x10^{-23}$
IL2RA 10p15.1	rs706779 A>G	0.467	0.85 (0.80-0.91)	$2.27x10^{-6}$	0.84 (0.74-0.96)	0.0125	0.858	0.85 (0.80-0.91)	$2.69x10^{-7}$
TSHR 14q31.1	rs2300519 T> *A*	0.380	1.54 (1.44-1.64)	$1.34x10^{-38}$	0.93 (0.81-1.07)	0.295	$4.64x10^{-12}$		

†Assuming 13 [8 (Table 1) and 5 (Table S1)] independent tests, the adjusted *P*-value was $3.85x10^{-3}$ for the 0.05 level of significance based on a Bonferroni correction for multiple testing.

Table S2: A summary of SNPs that show large allele frequency differences ($P < 1.12x10^{-6}$) between controls from the 12 geographical regions of Great Britain (Materials and Methods). Genotype signal intensity cluster plots have been examined visually for these highly differential SNPs (two SNPs were dropped rs2310187/2q12.1 and rs12210050/6p25.3). We note that the most significant *LCT* SNP, rs2236783/2q21.3 ($P = 1.41x10^{-5}$), was above the threshold *P*-value. Positions are in NCBI build 36 co-ordinates. distance in base pairs (dist).

Chromosome	Gene	Reported SNP	Position (base pairs)	P-value
2	IL1R2(dist=99131),IL1R1(dist=26387)	rs12470623	102110447	$1.52e-07$
2	IL1R2(dist=101821),IL1R1(dist=23697)	rs6752379	102113137	$1.52e-07$
2	IL1R2(dist=116654),IL1R1(dist=8864)	rs11123911	102127970	$1.51e-07$
2	IL1R2(dist=118953),IL1R1(dist=6565)	rs13035227	102130269	$1.04e-06$
2	IL1R2(dist=119894),IL1R1(dist=5624)	rs6754776	102131210	$1.09e-06$
2	IL1R1	rs871658	102138297	$1.07e-06$
2	RAB3GAP1	rs7570971	135554376	$9.01e-09$
2	RAB3GAP1	rs6730157	135623558	$4.32e-09$
2	RAB3GAP1(dist=26518),ZRANB3(dist=2777)	rs1375131	135671267	$2.22e-09$
2	ZRANB3	rs1561277	135808531	$3.49e-07$
2	R3HDM1	rs56369224	136045360	$5.29e-08$
2	R3HDM1	rs12465802	136097818	$1.10e-07$
2	MCM6	rs4988235	136325116	$2.54e-09$
2	MCM6	rs182549	136333224	$6.56e-09$
2	DARS	rs6754311	136424452	$4.49e-09$
2	DARS	rs12615624	136438073	$2.74e-07$
4	TLR1(dist=9090),TLR6(dist=9827)	rs4833103	38491897	$2.40e-10$
6	IRF4	rs12203592	341321	$4.31e-22$
6	IRF4(dist=9838),EXOC2(dist=63857)	rs62389423	366281	$1.44e-15$
6	IRF4(dist=11188),EXOC2(dist=62507)	rs62389424	367631	$3.69e-15$
6	IRF4(dist=13472),EXOC2(dist=60223)	rs9405661	369915	$1.09e-06$
6	IRF4(dist=20607),EXOC2(dist=53088)	rs6925797	377050	$1.52e-07$
16	ZNF764	imm_16_30477512	30477512	$1.67e-07$
16	ZNF785	rs9934806	30500560	$1.02e-07$
16	ZNF689	rs8063005	30525199	$9.15e-08$
16	PHKG2	rs72793380	30666244	$5.76e-07$

Supplementary Figures

Figure S1: Single-locus 1-df Cochran-Armitage trend test results at the recently reported novel AITD susceptibility locus 6q27 (1). The most disease associated SNP, imm−6−167338101 (P = $1.6x10^{-7}$ in HT and GD patients/controls; Table 1), was located within *FGFR10P*. The top panel — GD patients and controls; the middle panel — HT patients and controls; and, the bottom panel — GD and HT patients and controls.

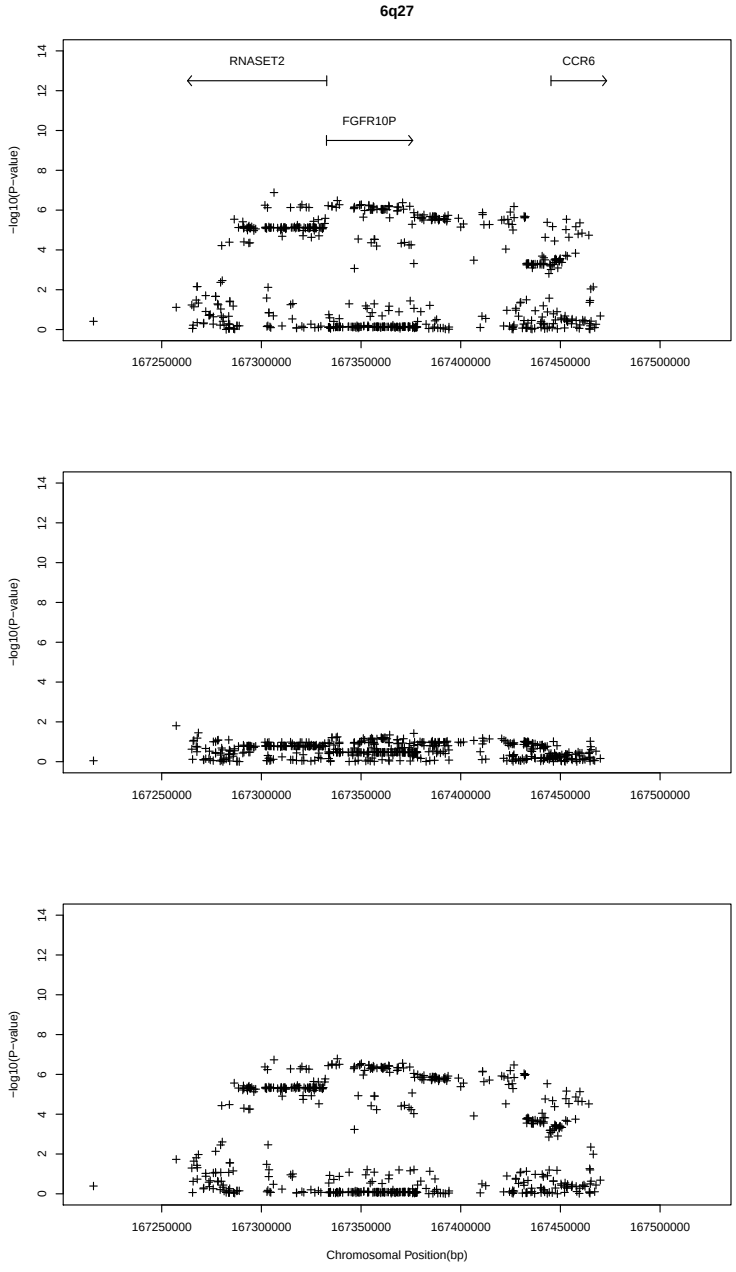


Figure S2: Single-locus 1-df Cochran-Armitage trend test results at the novel AITD susceptibility locus 1q36.32. The most disease associated SNP, rs2843403 ($P = 7.9x10^{-7}$ in GD patients/controls; Table 1), was located within *MMEL1*. The top panel — GD patients and controls; the middle panel — HT patients and controls; and, the bottom panel — GD and HT patients and controls.

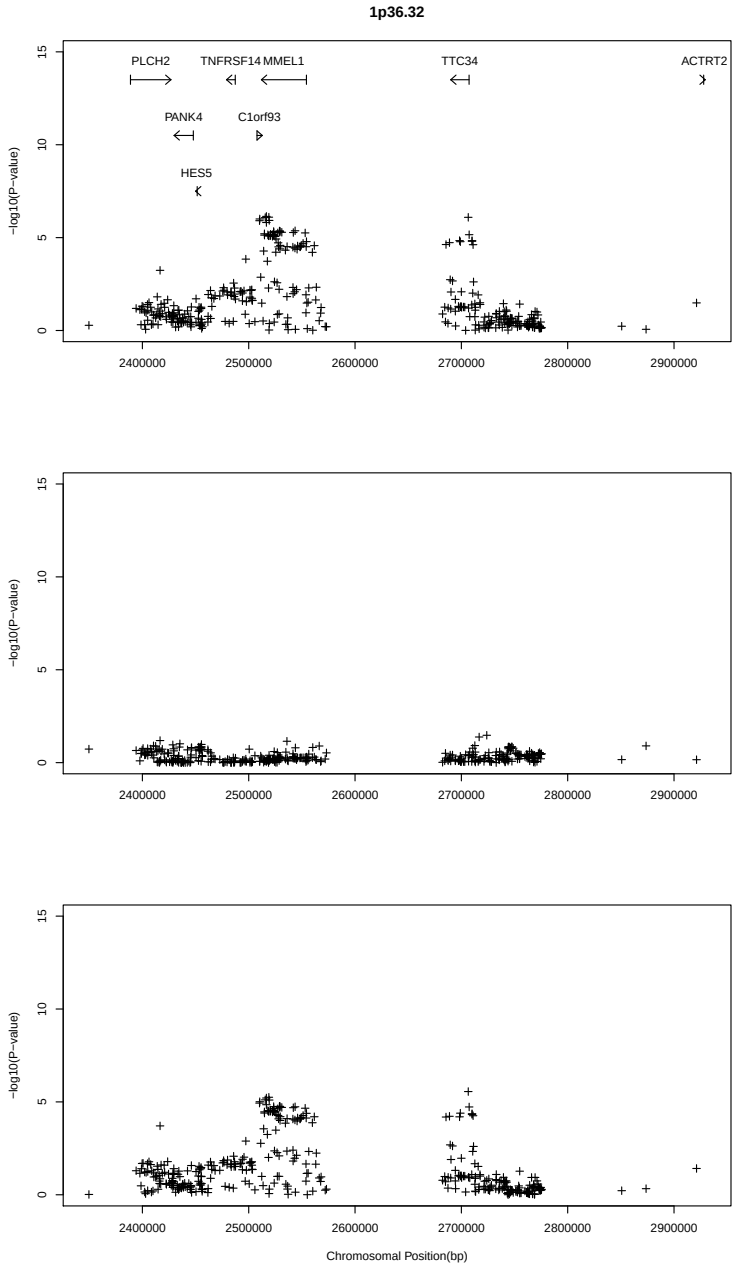


Figure S3: Single-locus 1-df Cochran-Armitage trend test results at the novel AITD susceptibility locus 3q27.3/3q28. The most disease associated SNP, rs13093110 (P = $3.7x10^{-8}$ in HT and GD patients/controls; Table 1), was located within *LPP*. The top panel — GD patients and controls; the middle panel — HT patients and controls; and, the bottom panel — GD and HT patients and controls.

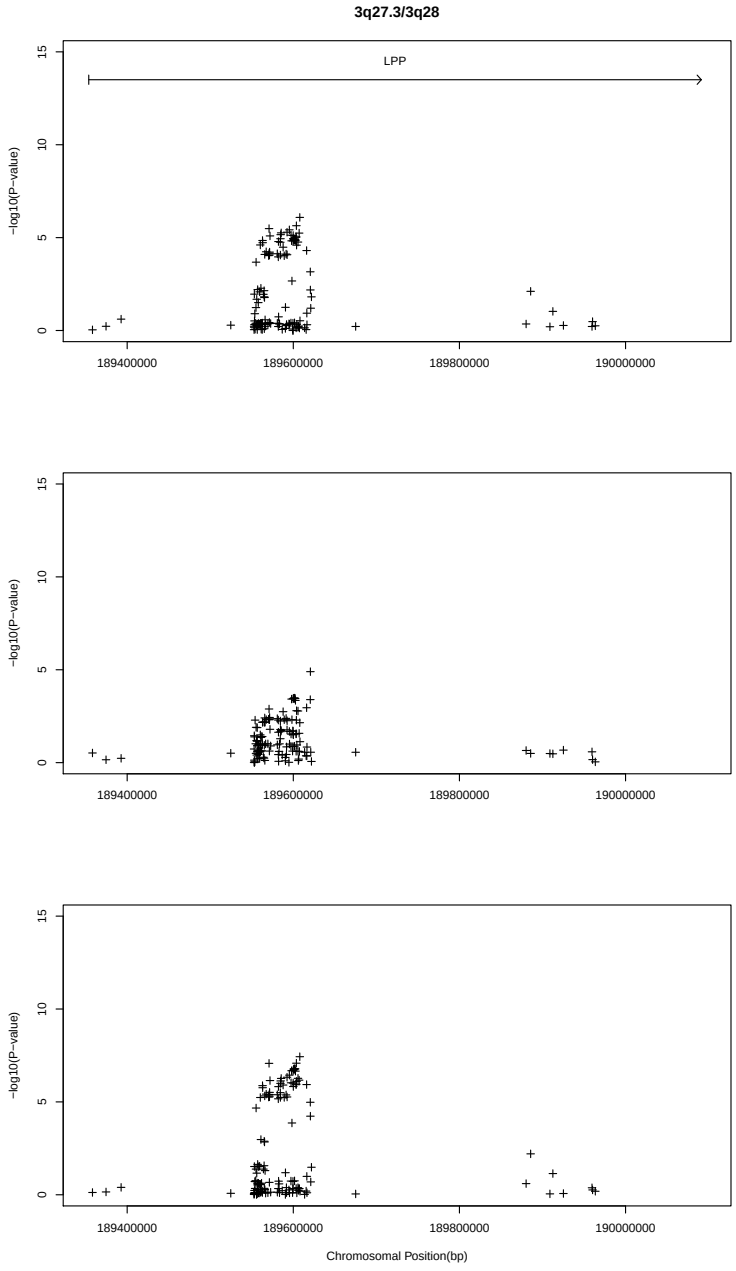


Figure S4: Single-locus 1-df Cochran-Armitage trend test results at the novel AITD susceptibility locus 6q15. The most disease associated SNP, rs72928038 (P = $1.2x10^{-7}$ in HT and GD patients/controls; Table 1), was located within *BACH2*. The top panel — GD patients and controls; the middle panel — HT patients and controls; and, the bottom panel — GD and HT patients and controls.

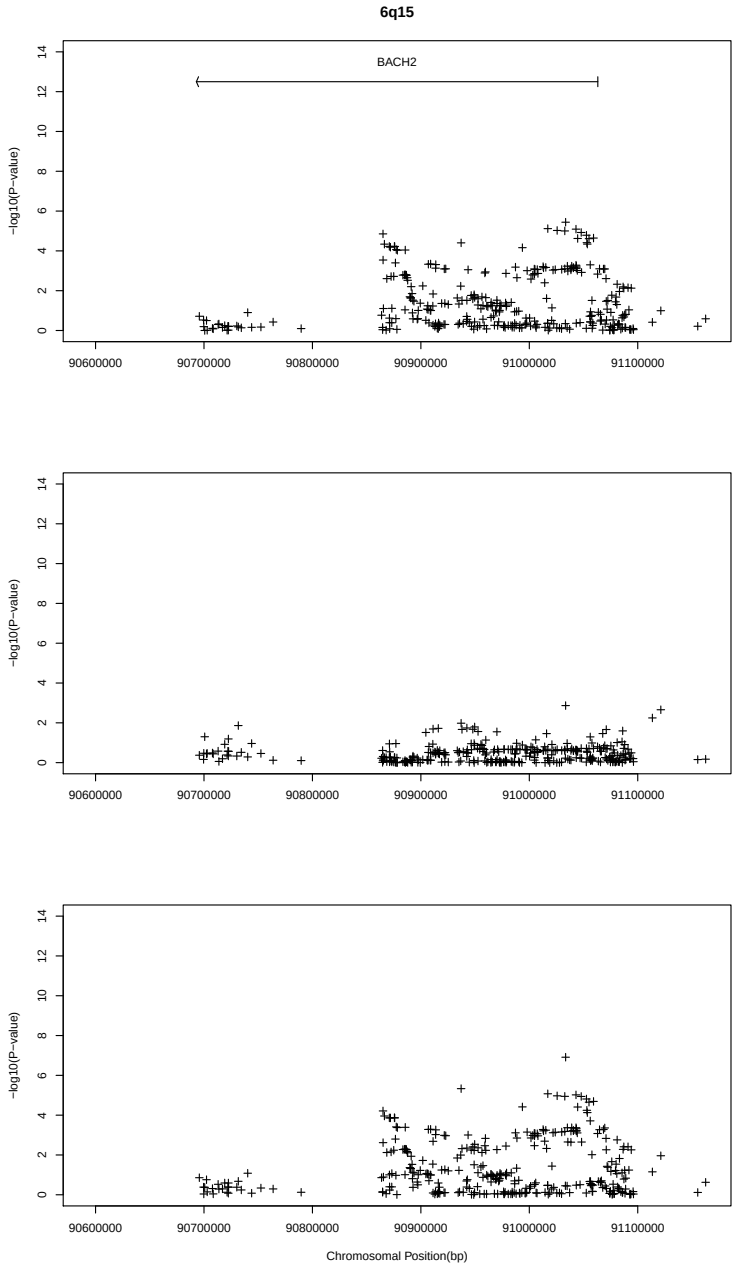


Figure S5: Single-locus 1-df Cochran-Armitage trend test results at the novel AITD susceptibility locus 12q12. The most disease associated SNP, rs57348955 ($P = 3.8x10^{-8}$ in GD patients/controls; Table 1), was located within *PRICKLE1*. The top panel — GD patients and controls; the middle panel — HT patients and controls; and, the bottom panel — GD and HT patients and controls.

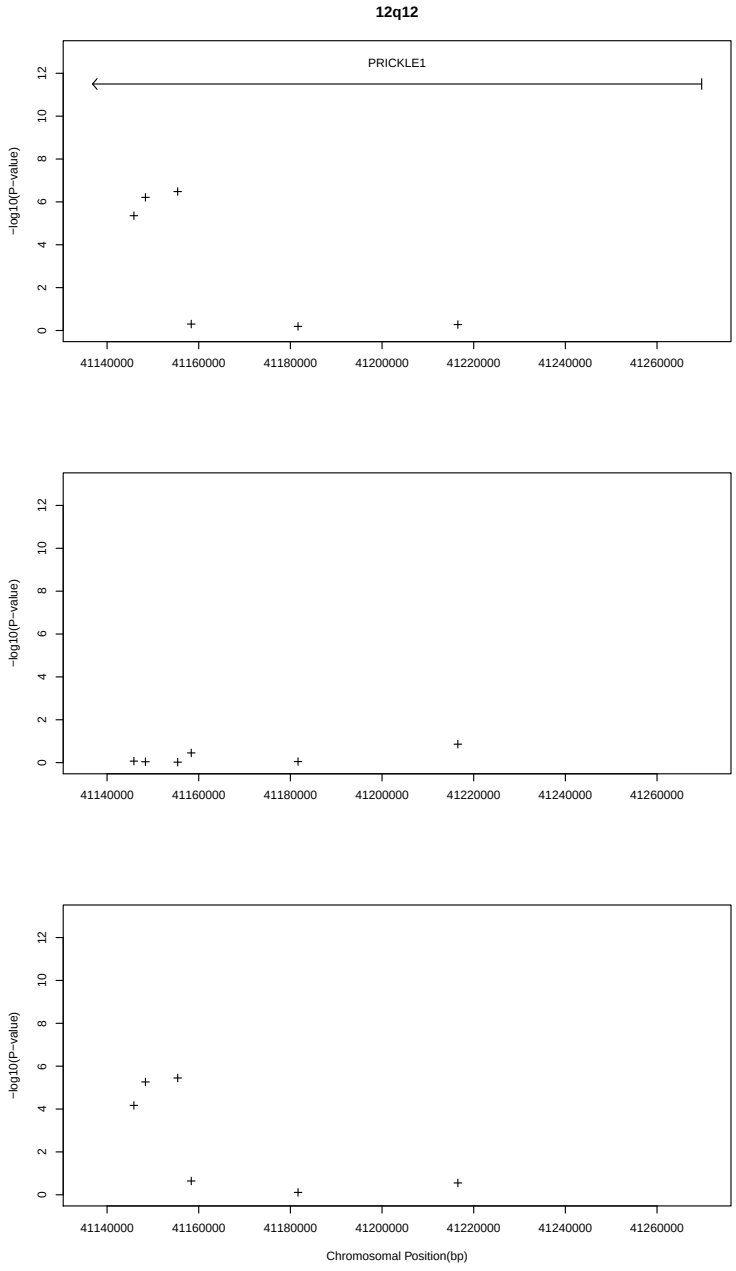


Figure S6: Single-locus 1-df Cochran-Armitage trend test results at the novel AITD susceptibility locus 16p11.2. The most disease associated SNP, rs57348955 ($P = 3.8x10^{-8}$ in GD patients/controls; Table 1), was located 82.6 kb upstream of *ITGAM*. The top panel — GD patients and controls; the middle panel — HT patients and controls; and, the bottom panel — GD and HT patients and controls.

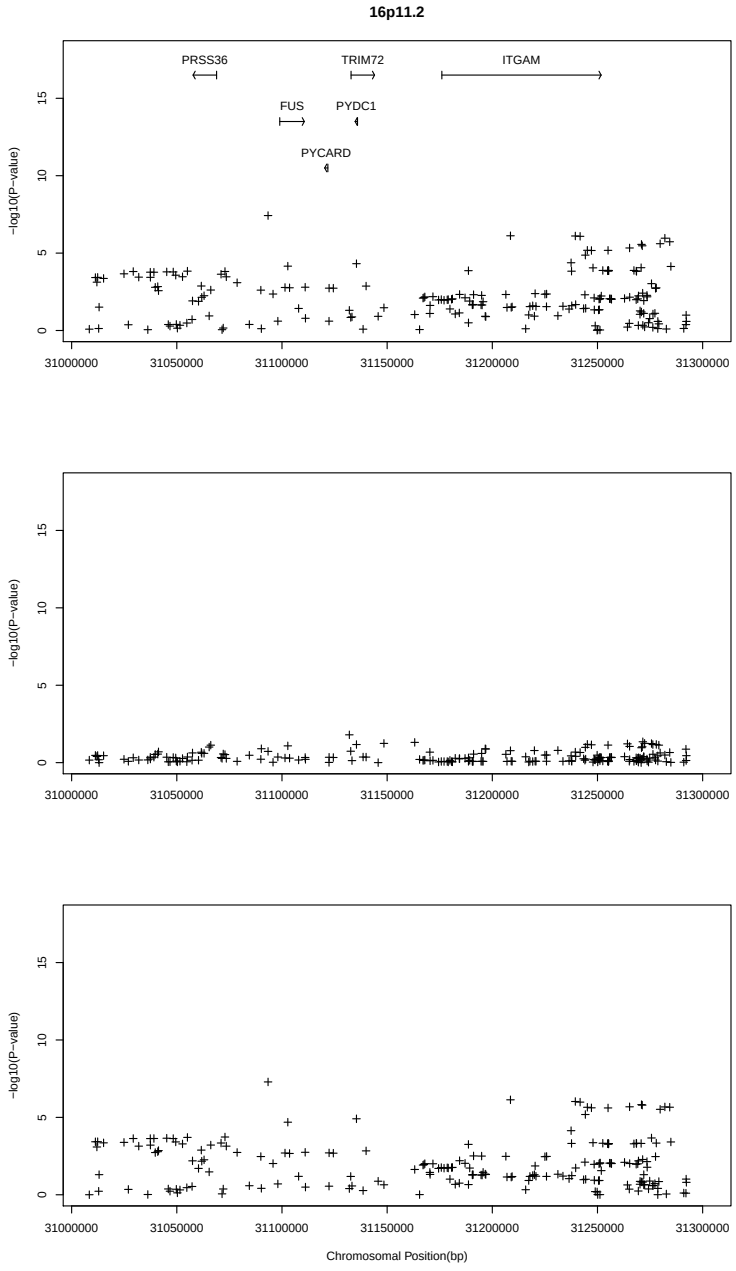


Figure S7: Single-locus 1-df Cochran-Armitage trend test results at the novel AITD susceptibility locus 2p25.1. The top panel — GD patients and controls; the middle panel — HT patients and controls; and, the bottom panel — GD and HT patients and controls.

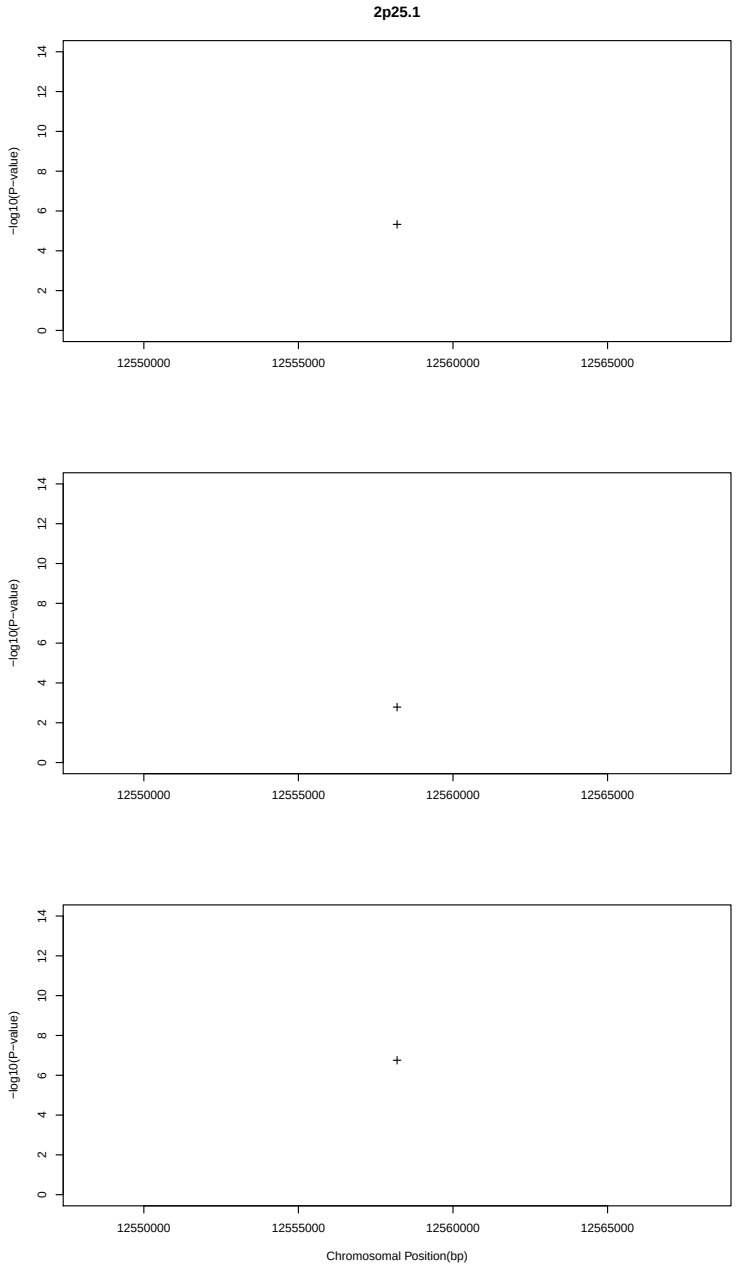


Figure S8: Single-locus 1-df Cochran-Armitage trend test results at the novel AITD susceptibility locus 11q21. The top panel — GD patients and controls; the middle panel — HT patients and controls; and, the bottom panel — GD and HT patients and controls.

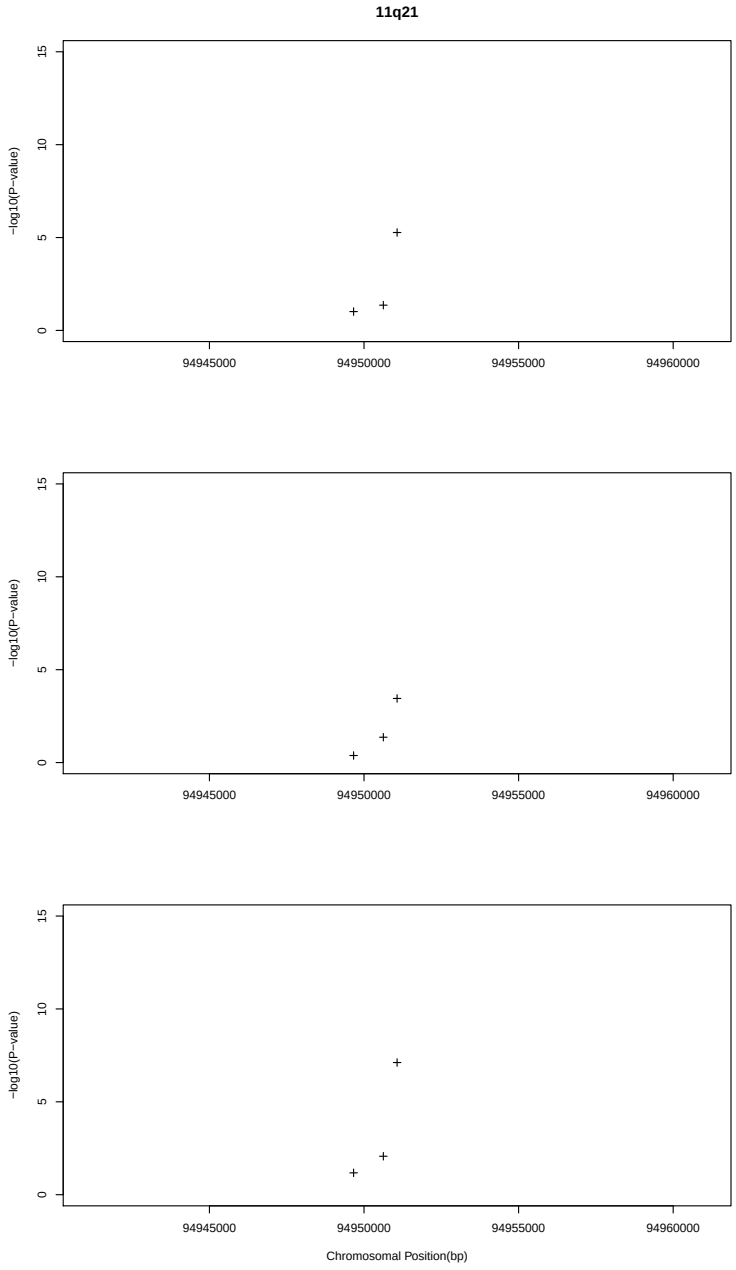


Figure S9: Per-sample call-rates versus heterozygosity across chromsomes 1 to 22 for WTSI control samples. Dashed lines denote the quality controls thresholds applied.

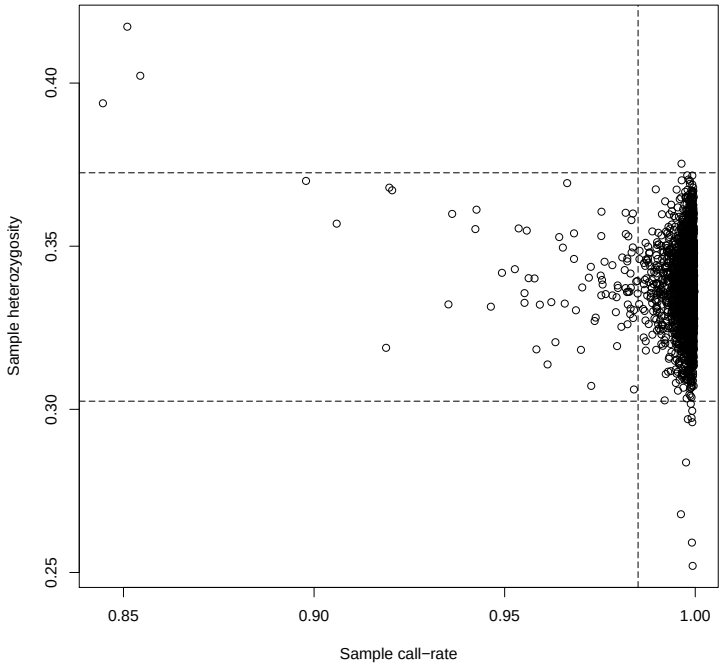


Figure S10: Per-sample call-rates versus heterozygosity across chromsomes 1 to 22 for UVA control samples. Dashed lines denote the quality controls thresholds applied.

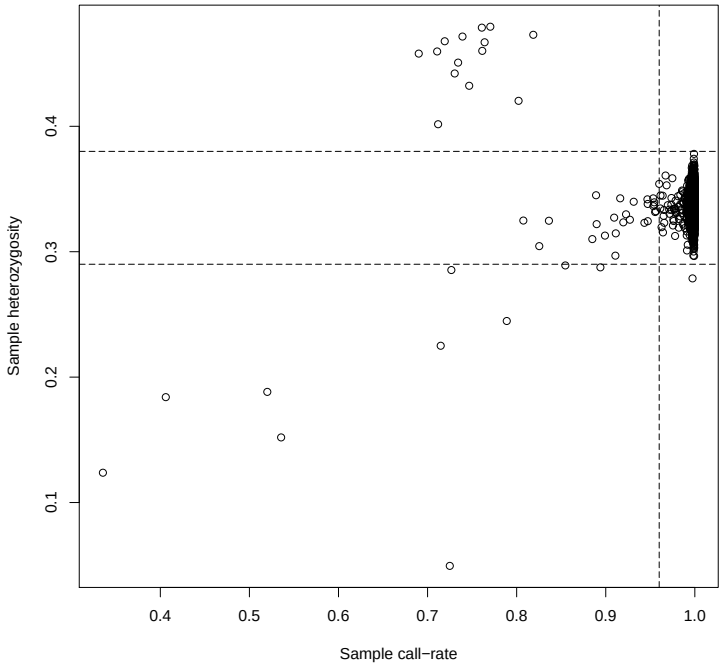


Figure S11: Per-sample call-rates versus heterozygosity across chromsomes 1 to 22 for AITD patient samples. Dashed lines denote the quality controls thresholds applied.

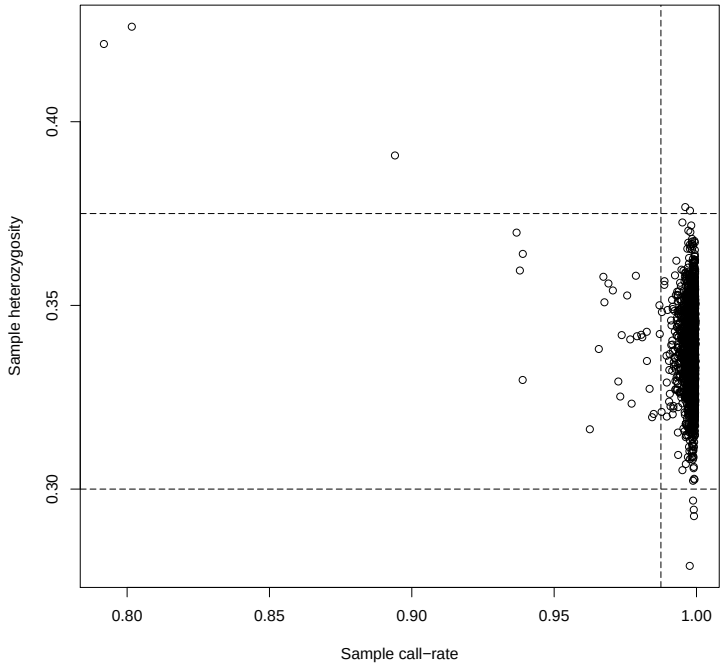


Figure S12: Ancestry clusters based on HapMap2 reference samples and SNPs available on the ImmunoChip that differentiated between the three HapMap populations (CEU, YRI and JPT+CHB) to derive two principal component scores for ancestry to exclude WTSI control subjects with substantial non-European ancestry. We excluded 18 WTSI controls subjects.

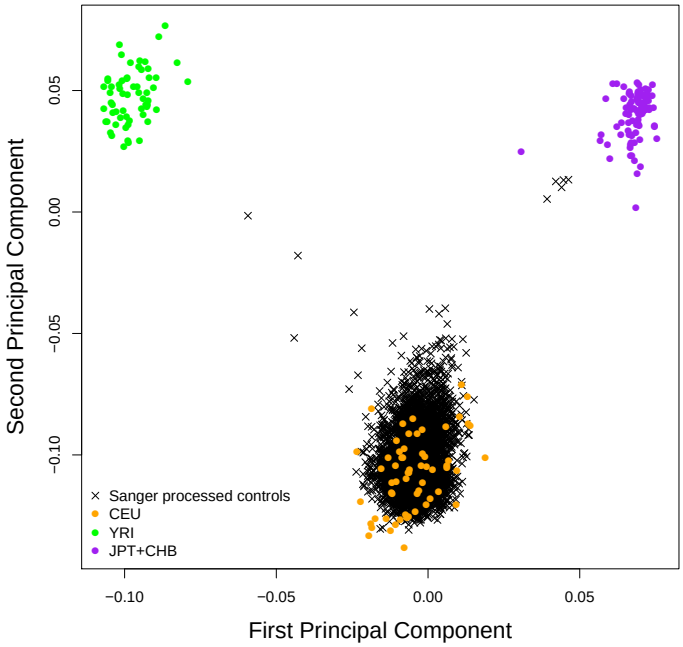


Figure S13: Ancestry clusters based on HapMap2 reference samples and SNPs available on the ImmunoChip that differentiated between the three HapMap populations (CEU, YRI and JPT+CHB) to derive two principal component scores for ancestry to exclude UVA control subjects with substantial non-European ancestry. We excluded five UVA control subjects.

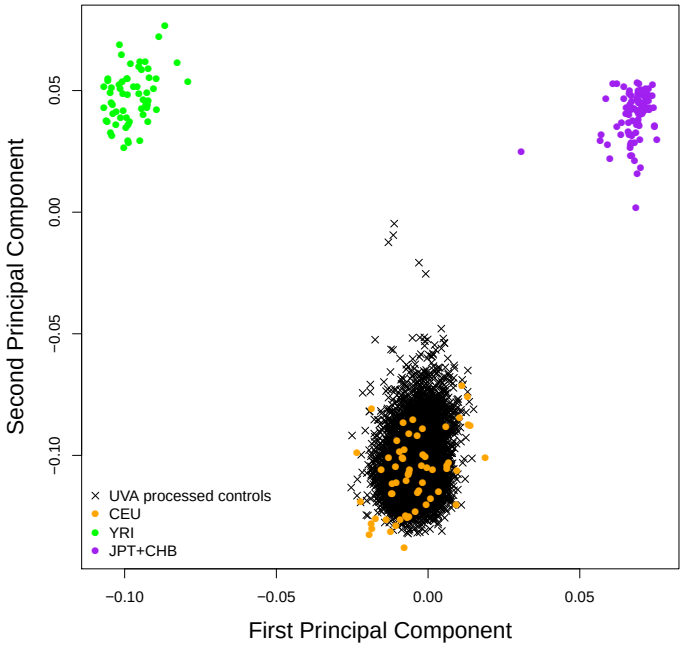


Figure S14: Ancestry clusters based on HapMap2 reference samples and SNPs available on the ImmunoChip that differentiated between the three HapMap populations (CEU, YRI and JPT+CHB) to derive two principal component scores for ancestry to exclude AITD patients with substantial non-European ancestry. We excluded three AITD patients.

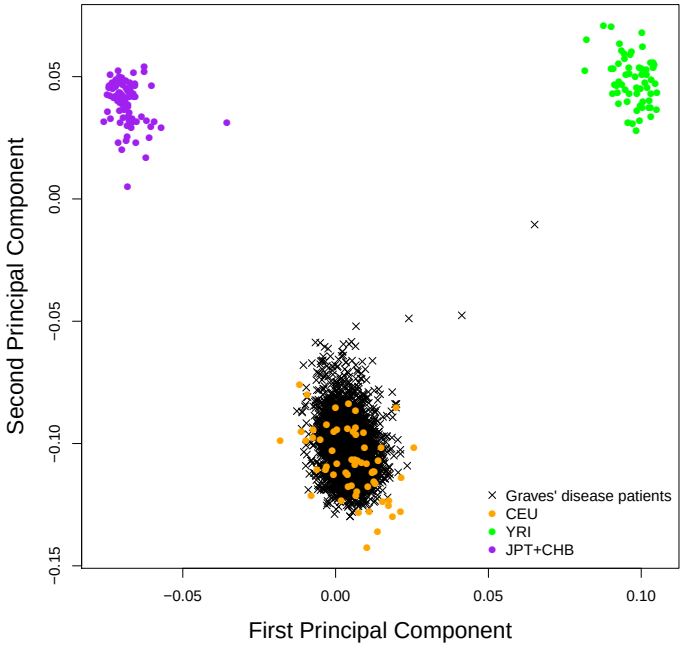
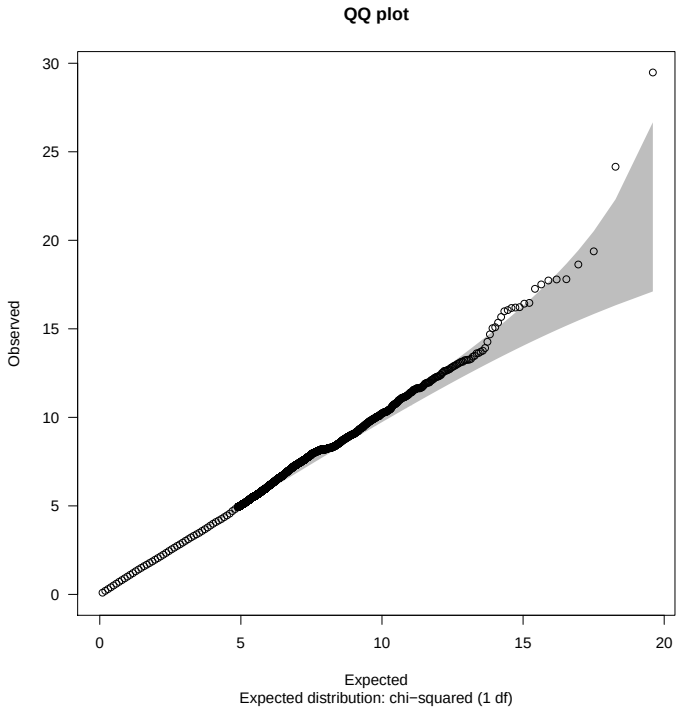


Figure S15: Quantile-quantile plot for the 1 degree-of-freedom association tests between WTSI and UVA control samples. The overdispersion factor (λ) of the 1-df association tests was 1.017.



References

1. Chu, X., *et al.* (2011) A genome-wide association study identifies two new risk loci for Graves' disease. *Nat Genet* **43**, 897-901.