

Published in final edited form as:

Nat Genet. ; 43(4): 329–332. doi:10.1038/ng.789.

Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis

George F. Mells^{1,2,*}, James A.B. Floyd^{3,*}, Katherine I. Morley^{3,4,*}, Heather J. Cordell⁵, Christopher S. Franklin³, So-Youn Shin³, Michael A. Heneghan⁶, James M. Neuberger⁷, Peter T. Donaldson⁸, Darren B. Day¹, Samantha J. Ducker⁸, Agnes W. Muriithi¹, Elizabeth F. Wheeler¹, Christopher J. Hammond⁹, Muhammad F. Dawwas², The UK PBC Consortium, The Wellcome Trust Case Control Consortium 3, David E. Jones⁸, Leena Peltonen³, Graeme J. Alexander², Richard N. Sandford^{1,#}, and Carl A. Anderson^{3,#}

¹ Academic Department of Medical Genetics, Cambridge University, Cambridge CB2 0QQ, UK

² Department of Hepatology, Cambridge University Hospitals NHS Foundation Trust, Cambridge CB2 0QQ, UK

³ Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1HH, UK

⁴ Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, School of Population Health, University of Melbourne, Australia

⁵ Institute of Human Genetics, Newcastle University, Newcastle upon Tyne NE1 3BZ, UK

⁶ Institute of Liver Studies, King's College Hospital NHS Foundation Trust, Denmark Hill, London SE5 9RS, UK

⁷ The Liver Unit, Queen Elizabeth Hospital, Birmingham B15 2TH, UK

Corresponding authors. .

*These authors contributed equally to this work

Author contributions: Study concept and design: GFM, HJC, MAH, JMN, PTD, The WTCCC3 management committee (JCB, CMB, DC, PJC, ADo, ADu, AH, CL, GL, APM, LM, NS, FOV), LP, DEJ, GJA, RNS, CAA.

PBC Genetics Study management: GFM, DBD, SJD, AWM, EFW, RNS.

Case ascertainment and phenotyping: GFM, DBD, SJD, AWM, EFW, MD, The UK PBC Consortium (AAI, YA, AAu, RA, SB, ABa, Abe, GB, PB, GBr, ABri, MB, ABro, JB, Abu, GBu, JBu, MCa, RCh, AC, CLC, NC, JC, MCo, MCr, RCr, HC, MCz, DD, MD, ADa, RD, SD, ADoud, ADoug, ADr, Red, CE, Rev, MF, NF, AF, JGa, IGe, KG, DGI, PG, IGo, DGo, FG, HG, JGo, AG, CG, RH, CHe, AH, BH, Cho, HH, SJ, KKa, MK, HK, SK, KKo, CL, SLA, SLev, AL, GLi, GLip, ML, NM, GM, TMac, AMa, SM, RMar, RMat, TMat, CMc, PMc, JMc, AMc, BM, JMe, PMi, CMi, HM, DN, TN, MN, AP, SPa, MPa, SPe, MPi, DP, MPr, RP, JRa, DRa, RR, IR, JRi, DRo, CR, SR, ASae, Ssa, IS, ASaw, JSa, SSe, ASh, TS, JSh, CS, RSh, VSW, US, CJS, ASi, Ssi, GS, PS, RSt, ATa, JTh, JTi, ATu, EU, GWa, MWe, RW, SWh, MWilk, EW, MWill, SWi, Gwo, MWr, TY, KY), DEJ, GA, RNS.

Control sample ascertainment: The UK Blood Service Controls group (JSa, JSt, WHO), The 1958 Birth Cohort Controls group (WLA, SMR, DPS), CJH, CAA.

Genotyping: WTCCC3 DNA, Genotyping and Informatics group (HB, SE, MG, EG, SEH, CL, SP, DS, PW).

Statistical Analysis: JABF, KIM, HJC, CSF, S-Ys, The WTCCC3 data analysis group (JCB, RM, NS, EZ), CAA.

Manuscript Preparation: GFM, JABF, HJC, DEJ, GJA, RNS, CAA.

Consortia: The UK PBC Consortium and The Wellcome Trust Case Control Consortium 3

A full list of members is available in the Supplementary Note

Competing financial interests: The authors declare no competing financial interests.

URLs used in this manuscript

WTCCC3: <http://www.WTCCC.org.uk>

OrageneTM: <http://www.dnagenotek.com>

Source Bioscience Healthcare: <http://healthcare.sourcebioscience.com>

twinsUK: <http://www.twinsuk.ac.uk>

Data access committee: <http://www.ebi.ac.uk/ega/>.

⁸ Institute of Cellular Medicine, Medical School, Newcastle University, Newcastle upon Tyne NE2 4HH, UK

⁹ Department of Twin Research & Genetic Epidemiology, King's College London SE1 7EH, UK

Abstract

In addition to the *HLA*-locus, six genetic risk factors for primary biliary cirrhosis (PBC) have been identified in recent genome-wide association studies (GWAS). To identify additional loci, we carried out a GWAS using 1,840 cases from the UK PBC Consortium and 5,163 UK population controls as part of the Wellcome Trust Case Control Consortium 3 (WTCCC3). Twenty-eight loci were followed up in an additional UK cohort of 620 PBC cases and 2,514 population controls. We identified 12 novel risk loci ($P < 5 \times 10^{-8}$) and replicated all previously associated loci. Three further novel loci were identified by meta-analysis of data from our study and previously published GWAS results. New candidate genes include *STAT4*, *DENND1B*, *CD80*, *IL7R*, *CXCR5*, *TNFRSF1A*, *CLEC16A*, and *NFKB1*. This study has considerably expanded our knowledge of the genetic architecture of PBC.

Primary biliary cirrhosis is a chronic, autoimmune liver disease characterized by non-suppurative destructive cholangitis and highly specific auto-antibodies to pyruvate dehydrogenase complex (PDC)¹. It is an important cause of chronic liver disease and a well-established indication for liver transplantation. In the UK, the prevalence of PBC is approximately 35 per 100,000 adults and 94 per 100,000 women aged 40 years or older². PBC has a sibling relative risk of ~10, suggesting a substantial genetic contribution to PBC³.

Previous studies have established that PBC is associated with *HLA-DR8*, with odds ratio ranging from 2.4 to 3.3 depending on the population examined⁴. To date, six non-*HLA* loci have been associated with PBC at a genome-wide level of significance ($P < 5 \times 10^{-8}$): *IL12A* (3q25)⁵, *IL12RB2* (1p31)⁵, *IRF5/TNPO3* (7q32)^{6,7}, *ORMDL3/IKZF3* (17q12)^{6,7}, *MMEL1* (1p36)⁶, and *SPIB* (19q13)⁷. To identify additional PBC risk loci we conducted a GWAS in a large cohort of UK PBC cases and population controls as part of the Wellcome Trust Case Control Consortium 3.

PBC cases were drawn from the UK PBC Consortium, which consists of 142 NHS Trusts in the UK, including all liver transplant centers. All cases were of self-declared British or Irish ancestry. PBC cases were genotyped using the Illumina 660W-Quad array. UK population controls were genotyped using the Illumina Human1M-Duo by the Wellcome Trust Case Control Consortium 2 and consisted of individuals from the 1958 British Birth Cohort and National Blood Service⁸. Following stringent quality control (see Online Methods), 507,467 SNPs were available across 1,840 cases and 5,163 'historical' population controls (see Supplementary Tables 1 and 2). The quantile-quantile plot of the case-control chi-square test statistics demonstrates a substantial excess of significant associations in the tail of the distribution, even after removal of known loci, which cannot be ascribed to overall inflation of the distribution (genomic control $\lambda = 1.09$; see Supplementary Figure 1).

We identified 34 loci where one or more SNPs showed at least suggestive evidence for association ($P < 1 \times 10^{-5}$), including six of the seven previously associated loci (Table 1). We found weaker evidence for replication at the seventh previously associated locus 1p36 (containing *MMEL1*; $P = 4 \times 10^{-3}$). 28 loci achieving at least suggestive significance and not previously associated with PBC at genome-wide significance were followed up by genotyping 46 SNPs in an independent panel of 620 cases from the UK PBC Consortium (Supplementary Table 3). A comparison set of 2,514 'historical' UK population controls previously genotyped by TwinsUK using the Illumina HumanHap610 array was obtained (see Online Methods). Twelve of these 28 loci were significant at $P < 0.05$ in the replication

study, and $P < 5 \times 10^{-8}$ in the combined analysis of the discovery and replication cohorts, and thus represent novel PBC associations (Table 2). Two of these loci (2q32 and 1q31) showed suggestive evidence of association in a previous PBC GWAS study⁵. Follow-up genotyping in large independent panels of cases and controls from a range of ethnicities is needed to further characterize these loci.

To identify additional risk loci, we combined summary statistics from our discovery cohort with those from the two datasets included in the previously published meta-analysis of PBC GWAS⁷. Three further novel loci reached genome-wide significance (see Table 3). This included one locus (14q32) that just failed to achieve genome-wide significance in our combined analysis of discovery and replication cohorts ($P = 1.69 \times 10^{-7}$), but did so with the addition of data from the study by Liu et al.⁷ ($P = 2.61 \times 10^{-13}$). As SNPs at the two other loci (3p24, 11q13) were not genotyped in our replication cohort, and the loci were identified based on summary statistics alone, genotyping using an independent technology in additional cohorts is needed to fully validate these associations. A combined GWAS meta-analysis is still warranted because we were only able to meta-analyze the top 100 SNPs from the Liu et al. study. Genome-wide imputation using HapMap3 reference panels did not identify any further genome-wide significant loci (see Online Methods, Supplementary Figure 2, and Supplementary Table 4), although for some loci imputed SNPs provided stronger evidence of association than the genotyped SNPs. No statistically significant gene-gene interactions were detected between associated loci, or after fitting an *HLA*-risk model (see Online Methods and Supplementary Table 5).

We found evidence for a second independent association at the 3q25 locus containing *IL12A* and *SCHIP1*, as did Liu et al.⁷ (see Supplementary Table 6). Three SNPs, located between *IL12A* and *SCHIP1*, remained genome-wide significant following a conditional logistic regression adjusting for the most significant SNP in the region (rs485499). Conducting the same analysis using the imputed data identified a further two SNPs reaching genome-wide significance. These five SNPs are all in linkage disequilibrium (LD; $r^2 > 0.2$) with each other, but none are in LD with rs485499. They are located downstream of *SCHIP1*, but upstream of *IL12A* (see Supplementary Figure 3). Fine mapping of this locus is needed to determine whether these association signals implicate independent variants affecting the same gene, or two different genes.

We identified plausible candidate genes within associated loci via manual curation, supported by evidence from: a) previous GWAS findings for other autoimmune diseases; b) GRAIL⁹, a literature-mining tool that identifies non-random, evidence-based links between genes; c) identification of non-synonymous SNPs in 1000 genomes data that are in LD ($r^2 > 0.8$) with the most associated genotyped SNP in each locus; d) identification of eQTL within associated loci that are in LD ($r^2 > 0.8$) with the most associated SNP at that locus, using data from Dixon et al.¹⁰ (see Online Methods and Supplementary Tables 7-10). Even in aggregate these analyses do not identify the gene(s) containing causal variants but they allow us to identify potential candidate genes for future follow-up studies. Supplementary Figure 4 shows all genes within each of the associated loci.

The results from the GWAS of PBC conducted to date provide additional support for the involvement of three pathways previously implicated in the pathogenesis of PBC: NF- κ B signaling, T-cell differentiation, and Toll-like receptor (TLR) and Tumor Necrosis Factor (TNF) signaling.

We identified several loci containing genes involved in activation of NF- κ B, a transcription factor which regulates expression of many genes involved in the immune response and is highly activated in other autoimmune disorders such as rheumatoid arthritis, multiple

sclerosis, and asthma¹¹. Its importance in PBC is suggested by evidence that NF- κ B modulates the balance of survival and apoptosis in activated hepatic stellate cells¹², and NF- κ B p50 $-/-$ mice show aggressive hepatic inflammation and fibrosis¹³. The locus we identified at 4q24 contains the *NFKB1* gene itself, and we identified genes in pathways leading to NF- κ B activation at four other loci: 22q13 (*TABI*), 12p13 (*TNFRSF1A*), 3q13 (*CD80*), and 11q13 (*RPS6KA4*).

Loci identified to date suggest a role for T-lymphocyte differentiation in the development of PBC. T_H1 immune responses have been implicated in many autoimmune diseases¹⁴ and may be involved in development of autoreactive T-cells, consistent with the putative role of PDC-specific autoreactive T_H1 cells in the pathogenesis of human PBC and animal disease models¹⁵. IL-12 signaling promotes T_H1-type immune responses by driving differentiation of activated, naïve T-cells to T_H1 cells¹⁶ and three loci containing genes involved in IL-12 signaling have been identified for PBC: 3q25 (*IL12A*) and 1p31 (*IL12RB2*) by Hirschfield et al.⁵, and 2q32 (*STAT4*) in this study. These results provide further support for the T_H1 hypothesis regarding PBC development.

Activation of TLR signaling, and its downstream effectors such as TNF α , is well described in PBC¹⁷. The 7q32 locus, identified by Liu et al.⁷, contains *IRF5* which is activated in response to TLR-signaling and leads to selective expression of TNF α . We identified a locus at 11q13 containing *RPS6KA4*, which suppresses TLR-dependent cytokine production¹⁸. TNF α is an activating factor for a number of intracellular pathways that determine the fate of hepatocytes, and thus plays a key role in liver homeostasis¹⁹. We identified three loci containing genes in TNF α signaling pathways: 12p13 (*TNFRSF1A*), 1q31 (*DENND1B*), and 14q32 (*TNFAIP2*). *TNFRSF1A* is one of two receptors for TNF α , and *TNFRSF1A* $-/-$ mice show attenuated liver fibrosis when compared to wild-type mice after administration of a potent hepatotoxin²⁰. *DENND1B* interacts directly with *TNFRSF1A*²¹ and has previously been associated with asthma²². TNF α signaling also directly induces *TNFAIP2* expression²³.

In summary, this is the first report in a new series of GWAS undertaken by the WTCCC3. Twelve novel PBC risk loci have been identified in this study of >7,000 European samples, making this the largest GWAS of PBC to date. In addition, a further three loci achieve genome-wide significance following meta-analysis with published data. For many of the associated loci we have identified plausible candidate genes that support the involvement of the innate and adaptive immune systems in PBC etiology, particularly signaling via the NF- κ B, TLR, and TNF pathways, although these findings require confirmation through fine-mapping, gene-expression and functional studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The PBC sample collection was funded by the Isaac Newton Trust, the PBC Foundation, and the Wellcome Trust (085925/Z/08/Z). The PBC Genetics Study is a portfolio study of the National Institute for Health Research Comprehensive Clinical Research Network (NIHR CRN, portfolio reference 5630). The WTCCC3 project is also supported by the Wellcome Trust (WT090355/A/09/Z, WT090355/B/09/Z). GFM is a Clinical Research Training Fellow of the Medical Research Council (G0800460). He is also supported by a Raymond and Beverly Sackler Studentship. CJH is an NIHR Senior Research Fellow. CAA is funded by the Wellcome Trust (WT91745/Z/10/Z).

We are grateful to The PBC Foundation for helping us to establish the PBC Genetics Study, for endorsing it, and for encouraging members of the Foundation to contribute samples. We thank all of the research nurses who have assisted with participant recruitment in collaborating centres. We thank staff in the NIHR CRN and CRC Cymru for

providing invaluable support. We are grateful to Kevin Chittock and his colleagues at Source Bioscience for their efforts. We thank Oliver Burren for designing the participant database and for providing IT support. We acknowledge the Department of Twin Research & Genetic Epidemiology, King's College London, for additional control genotype data, which was part-funded by a US NIH/NEI grant 1R01EY018246 (PI Terri Young), and genotyping by the NIH CIDR, with additional genotyping performed at the WTSI. The study also receives support from the NIHR comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust partnering KCL. Finally, we thank the individuals who contributed samples used in this study.

References

1. Kaplan MM, Gershwin ME. Primary biliary cirrhosis. *N Engl J Med*. 2005; 353:1261–73. [PubMed: 16177252]
2. James OF, et al. Primary biliary cirrhosis once rare, now common in the United Kingdom? *Hepatology*. 1999; 30:390–4. [PubMed: 10421645]
3. Jones DE, Watt FE, Metcalf JV, Bassendine MF, James OF. Familial primary biliary cirrhosis reassessed: a geographically-based population study. *J Hepatol*. 1999; 30:402–7. [PubMed: 10190721]
4. Donaldson PT, et al. HLA class II alleles, genotypes, haplotypes, and amino acids in primary biliary cirrhosis: a large-scale study. *Hepatology*. 2006; 44:667–74. [PubMed: 16941709]
5. Hirschfield GM, et al. Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants. *N Engl J Med*. 2009; 360:2544–55. [PubMed: 19458352]
6. Hirschfield GM, et al. Variants at IRF5-TNPO3, 17q12-21 and MMEL1 are associated with primary biliary cirrhosis. *Nat Genet*. 2010; 42:655–7. [PubMed: 20639879]
7. Liu X, et al. Genome-wide meta-analyses identify three loci associated with primary biliary cirrhosis. *Nat Genet*. 2010; 42:658–60. [PubMed: 20639880]
8. Barrett JC, et al. Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nat Genet*. 2009; 41:1330–4. [PubMed: 19915572]
9. Raychaudhuri S, et al. Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genet*. 2009; 5:e1000534. [PubMed: 19557189]
10. Dixon AL, et al. A genome-wide association study of global gene expression. *Nat Genet*. 2007; 39:1202–7. [PubMed: 17873877]
11. Li Q, Verma IM. NF-kappaB regulation in the immune system. *Nat Rev Immunol*. 2002; 2:725–34. [PubMed: 12360211]
12. Elsharkawy AM, et al. The NF-kappaB p50:p50:HDAC-1 repressor complex orchestrates transcriptional inhibition of multiple pro-inflammatory genes. *J Hepatol*. 2010
13. Price AL, et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006; 38:904–9. [PubMed: 16862161]
14. Zhernakova A, van Diemen CC, Wijmenga C. Detecting shared pathogenesis from the shared genetics of immune-related diseases. *Nat Rev Genet*. 2009; 10:43–55. [PubMed: 19092835]
15. Jones DE. Pathogenesis of primary biliary cirrhosis. *Gut*. 2007; 56:1615–24. [PubMed: 17641080]
16. Trinchieri G. Interleukin-12 and the regulation of innate resistance and adaptive immunity. *Nat Rev Immunol*. 2003; 3:133–46. [PubMed: 12563297]
17. Mao TK, et al. Altered monocyte responses to defined TLR ligands in patients with primary biliary cirrhosis. *Hepatology*. 2005; 42:802–8. [PubMed: 16175622]
18. Vermeulen L, Berghe WV, Beck IM, De Bosscher K, Haegeman G. The versatile role of MSKs in transcriptional regulation. *Trends Biochem Sci*. 2009; 34:311–8. [PubMed: 19464896]
19. Tacke F, Luedde T, Trautwein C. Inflammatory pathways in liver homeostasis and liver injury. *Clin Rev Allergy Immunol*. 2009; 36:4–12. [PubMed: 18600481]
20. Kitamura K, et al. Pathogenic roles of tumor necrosis factor receptor p55-mediated signals in dimethylnitrosamine-induced murine liver fibrosis. *Lab Invest*. 2002; 82:571–83. [PubMed: 12003998]
21. Del Villar K, Miller CA. Down-regulation of DENN/MADD, a TNF receptor binding protein, correlates with neuronal cell death in Alzheimer's disease brain and hippocampal neurons. *Proc Natl Acad Sci U S A*. 2004; 101:4210–5. [PubMed: 15007167]

22. Sleiman PM, et al. Variants of DENND1B associated with asthma in children. *N Engl J Med.* 2010; 362:36–44. [PubMed: 20032318]
23. Sarma V, Wolf FW, Marks RM, Shows TB, Dixit VM. Cloning of a novel tumor necrosis factor- α -inducible primary response gene that is differentially expressed in development and capillary tube-like formation in vitro. *J Immunol.* 1992; 148:3302–12. [PubMed: 1374453]
24. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* 2007; 81:559–75. [PubMed: 17701901]
25. Morris JA, Randall JC, Maller JB, Barrett JC. Evoker: a visualization tool for genotype intensity data. *Bioinformatics.* 2010; 26:1786–7. [PubMed: 20507892]
26. Magi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics.* 2010; 11:288. [PubMed: 20509871]
27. Browning BL, Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am J Hum Genet.* 2009; 84:210–23. [PubMed: 19200528]
28. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria:

Table 1

Association results and *in silico* analyses for the 7 previously confirmed primary biliary cirrhosis risk loci.

CHR	SNP	RISK ALLELE	LEFT/RIGHT REGION (MB)	CANDIDATE GENE ^d (Number of genes in region)	A1 ^b	rsSNP ^c	GIWAS Cohort			
							Control RAF ^d	Case RAF ^d	P-value	OR (95% CI)
1p36	rs10752747	T	2.39 - 2.78	<i>MMEI1</i> (6)	○		0.339	0.367	2.65×10 ⁻³	1.13 (1.04 - 1.22)
1p31	rs17129789	C	67.53 - 67.71	<i>IL12RB2</i> (2)	○		0.177	0.247	9.48×10 ⁻²⁰	1.52 (1.39 - 1.67)
3q25	rs485499	T	160.96 - 161.3	<i>IL12A</i> (2)	●		0.574	0.651	2.29×10 ⁻¹⁶	1.38 (1.28 - 1.50)
6p21	rs7774434	C	26.21 - 33.74	Many (<i>MHC</i>)	○	●	0.379	0.494	3.86×10 ⁻³⁴	1.60 (1.48 - 1.73)
7q32	rs12531711	G	128.33 - 128.57	<i>IRF5</i> (2)	●		0.107	0.159	8.90×10 ⁻¹⁷	1.58 (1.41 - 1.76)
17q12	rs7208487	T	34.61 - 35.49	<i>ORMDL3</i> (23)	●		0.840	0.874	7.89×10 ⁻⁷	1.32 (1.18 - 1.48)
19q13	rs3745516	A	55.52 - 55.73	<i>SPIB</i> (11)		●	0.226	0.287	1.63×10 ⁻¹³	1.38 (1.32 - 1.44)

^aThe putative candidate gene represents the strongest candidate within the region based on available evidence, but does not preclude the existence of other plausible candidate genes within the region. The number of genes is based upon the RefSeq gene track.

^bOpen circles indicate the locus has been previously associated with another autoimmune disease; a filled circle indicates the same candidate gene has also been suggested.

^cFilled circles indicate a non-synonymous SNP in LD with our top SNP was identified in the candidate gene.

^dRAF indicates the risk allele frequency.

Table 2

Association results and *in silico* analyses for 12 newly confirmed primary biliary cirrhosis risk loci.

CHR	SNP	RISK ALLELE	LEFT/RIGHT REGION (MB)	CANDIDATE GENE ^a (Number of genes in region)	A1 ^b	GRAIL ^c	nsSNP ^d	GWAS cohort			Replication cohort			Combined samples			
								Control RAF ^e	Case RAF ^e	P-value	OR (95% CI)	Control RAF ^e	Case RAF ^e	P-value	OR (95% CI)	P-value	OR (95% CI)
1q31	rs12134279	T	195.58 - 196.21	<i>DENND1B</i> (4)	○	●		0.202	0.250	1.07×10 ⁻⁹	1.32 (1.21 - 1.44)	0.201	0.263	2.57×10 ⁻⁶	1.42 (1.23 - 1.65)	2.06×10 ⁻¹⁴	1.34 (1.23 - 1.45)
2q32	rs10931468	A	190.77 - 191.61	<i>STAF7</i> (7)	●	●		0.119	0.164	2.55×10 ⁻¹²	1.46 (1.31 - 1.62)	0.120	0.183	2.64×10 ⁻⁹	1.64 (1.39 - 1.94)	2.35×10 ⁻¹⁹	1.50 (1.37 - 1.64)
3q13	rs2293370	G	120.58 - 120.79	<i>CD80</i> (6)	●	●		0.804	0.853	7.70×10 ⁻¹¹	1.41 (1.27 - 1.56)	0.809	0.835	0.036	1.19 (1.01 - 1.41)	2.53×10 ⁻¹¹	1.35 (1.23 - 1.47)
4q24	rs7665090	C	103.61 - 104.24	<i>NFKB1</i> (7)	●	●		0.524	0.572	5.33×10 ⁻⁷	1.21 (1.13 - 1.31)	0.513	0.593	5.50×10 ⁻⁷	1.38 (1.22 - 1.57)	4.06×10 ⁻¹²	1.26 (1.18 - 1.34)
5p13	rs860413	A	35.74 - 36.08	<i>IL7R</i> (5)	●	●	●	0.719	0.773	3.09×10 ⁻¹⁰	1.33 (1.22 - 1.45)	0.729	0.769	4.50×10 ⁻³	1.24 (1.07 - 1.43)	1.02×10 ⁻¹¹	1.30 (1.21 - 1.40)
7p14	rs6974491	A	37.32 - 37.41	(0)				0.170	0.205	3.39×10 ⁻⁶	1.25 (1.14 - 1.38)	0.177	0.215	2.40×10 ⁻³	1.27 (1.09 - 1.49)	4.44×10 ⁻⁸	1.25 (1.16 - 1.36)
11q23	rs6421571	C	117.82 - 118.30	<i>CXCR5</i> (10)	○	●		0.809	0.855	3.53×10 ⁻¹⁰	1.40 (1.26 - 1.55)	0.810	0.847	2.10×10 ⁻³	1.30 (1.10 - 1.55)	2.69×10 ⁻¹²	1.37 (1.25 - 1.50)
12p13	rs1800693	C	6.29 - 6.33	<i>TNFRSF1A</i> (3)	●	●		0.401	0.452	5.51×10 ⁻⁸	1.23 (1.14 - 1.33)	0.403	0.445	8.70×10 ⁻³	1.18 (1.04 - 1.34)	1.80×10 ⁻⁹	1.22 (1.14 - 1.30)
14q24	rs911263	T	67.34 - 67.98	<i>RAD51L1</i> (2)	●	●		0.712	0.764	1.68×10 ⁻⁹	1.31 (1.20 - 1.43)	0.717	0.760	2.30×10 ⁻³	1.25 (1.08 - 1.45)	1.76×10 ⁻¹¹	1.29 (1.20 - 1.39)
16p13	rs12924729	G	10.92 - 11.22	<i>CLEC16A</i> (3)	●	●		0.679	0.737	7.68×10 ⁻¹¹	1.32 (1.21 - 1.44)	0.680	0.718	8.80×10 ⁻³	1.20 (1.05 - 1.38)	2.95×10 ⁻¹²	1.29 (1.20 - 1.38)
16q24	rs1117432	G	84.55 - 84.58	(0)	○			0.760	0.808	1.20×10 ⁻⁶	1.26 (1.15 - 1.39)	0.774	0.838	9.52×10 ⁻⁷	1.52 (1.28 - 1.79)	4.66×10 ⁻¹¹	1.31 (1.21 - 1.43)
22q13	rs968451	T	37.87 - 38.19	<i>MAP3K7IP1</i> (3)				0.194	0.233	4.31×10 ⁻⁷	1.27 (1.16 - 1.39)	0.193	0.237	6.45×10 ⁻⁴	1.30 (1.12 - 1.51)	1.08×10 ⁻⁹	1.27 (1.18 - 1.38)

PBC loci that meet genome-wide significance $P < 5 \times 10^{-8}$ in the combined analysis and $P < 0.05$ in the replication cohort. GWAS and replication cohort data for the replicated SNPs were merged using PLINK.

^aThe putative candidate gene represents the strongest candidate within the region based on available evidence, but does not preclude the existence of other plausible candidate genes within the region. The number of genes is based upon the RefSeq gene track.

^bOpen circles indicate the locus as been previously associated with another autoimmune disease; a filled circle indicates the same candidate gene has also been suggested.

^cFilled circles indicate the gene was identified by GRAIL as the most plausible functional candidate in the region ($P_{\text{text}} < 0.01$). GRAIL results are not provided for previously confirmed loci as these were used as seeds in the analysis.

^dFilled circles indicate a non-synonymous SNP in LD ($r^2 > 0.8$) with our top SNP was identified in the candidate gene.

^eRAF indicates the risk allele frequency.

Table 3
Genomic regions reaching genome-wide significance after meta-analysis with Liu et al. (2010) data.

CHR	SNP	RISK ALLELE	LEFT/RIGHT REGION (MB)	CANDIDATE GENE ^a (Number of genes in region)	Discovery sample			Liu et al. 2010			Meta-analysis		
					Control RAF ^b	Case RAF ^b	P-value	OR (95% CI)	P-value	OR ^c	P-value	OR (95% CI)	
3p24	rs1372072	A	16.82 - 17.13	<i>PLCL2</i> (1)	0.365	0.400	1.38×10 ⁻⁴	1.16 (1.08 - 1.25)	1.52×10 ⁻⁵	1.27	2.28×10 ⁻⁸	1.20 (1.12 - 1.27)	
11q13	rs538147	G	63.60 - 64.04	<i>RPS6KA4</i> (20)	0.606	0.647	1.01×10 ⁻⁵	1.19 (1.10 - 1.29)	7.72×10 ⁻⁶	1.28	2.06×10 ⁻¹⁰	1.23 (1.15 - 1.31)	
14q32	rs8017161	A	102.54 - 102.68	<i>TNFAIP2</i> (3)	0.396	0.439	4.71×10 ⁻⁶	1.20 (1.11 - 1.29)	4.86×10 ⁻⁷	1.31	2.61×10 ⁻¹³	1.22 (1.16 - 1.27)	

^aThe putative candidate gene represents the strongest candidate within the region based on available evidence, but does not preclude the existence of other plausible candidate genes within the region. The number of genes is based upon the RefSeq gene track.

^bRAF indicates the risk allele frequency.

^cLiu et al. (2010) do not provide confidence intervals for the odds ratios (ORs) estimated from their meta-analysis for these SNPs. None of these genomic regions have previously been associated at genome-wide significance with another autoimmune disease. GRAIL failed to identify any strong candidate genes within these regions and no nsSNPs were identified in high LD ($r^2 > 0.8$) with the most associated SNP at each locus.