Fine mapping and replication of genetic risk loci in primary sclerosing cholangitis

BRIJESH SRIVASTAVA¹, GEORGE F MELLS¹, HEATHER J CORDELL², AGNES MURIITHI¹, MATTHEW BROWN¹, EVA ELLINGHAUS³, ANDRE FRANKE³, UK-PSC CONSORTIUM, TOM H KARLSEN⁴,⁵,⁶, RICHARD N SANDFORD¹, GRAEME J ALEXANDER⁷, ROGER W CHAPMAN⁸, SIMON M RUSHBROOK⁹ & ESPEN MELUM¹⁴,⁵,⁶

¹Academic Department of Medical Genetics, University of Cambridge, Cambridge, UK, ²Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK, ³Institute of Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Germany, ⁴Norwegian PSC Research Center, Division of Cancer, Surgery and Transplantation, Oslo University Hospital, Rikshospitalet, Oslo, Norway, ⁵Research Institute for Internal Medicine, Division of Cancer, Surgery and Transplantation, Oslo University Hospital, Rikshospitalet, Oslo, Norway, ⁶Institute of Clinical Medicine, University of Oslo, Oslo, Norway, ⁷Department of Medicine, Division of Hepatology, University of Cambridge, Cambridge, UK, ⁸Department of Hepatology, John Radcliffe University Hospitals NHS Trust, Oxford, UK, and ⁹Department of Hepatology, Norfolk and Norwich University Hospitals NHS Trust, Norwich, UK

Abstract

Background and aims. Primary sclerosing cholangitis (PSC) is a chronic cholestatic liver disease characterized by progressive inflammation and fibrosis of the bile ducts eventually leading to biliary cirrhosis. Recent genetic studies in PSC have identified associations at 2q13, 2q35, 3p21, 4q27, 13q31 and suggestive association at 10p15. The aim of this study was to further characterize and refine the genetic architecture of PSC. Methods. We analyzed previously reported associated SNPs at four of these non-HLA loci and 59 SNPs tagging the IL-2/IL-21 and IL2RA loci in 992 UK PSC cases and 5162 healthy UK controls. Results. The most associated SNPs identified were rs3197999 (3p21 (MST1), p = 1.9 × 10⁻⁶, OR_A vs G = 1.28, 95% CI (1.16–1.42)); rs4147359 (10p15 (IL2RA), p = 2.6 × 10⁻⁴, OR_A vs G = 1.20, 95% CI (1.09–1.33)) and rs12511287 (4q27 (IL-2/IL-21), p = 3.0 × 10⁻⁴, OR_A vs T = 1.21, 95% CI (1.09–1.35)). In addition, we performed a meta-analysis for selected SNPs using published summary statistics from recent studies. We observed genome-wide significance for rs3197999 (3p21 (MST1), Pcombined = 3.8 × 10⁻¹²) and rs4147359 (10p15 (IL2RA), Pcombined = 1.5 × 10⁻⁸). Conclusion. We have for the first time confirmed the association of PSC with genetic variants at 10p15 (IL2RA) locus at genome-wide significance and replicated the associations at MST1 and IL-2/IL-21 loci in a large homogeneous UK population. These results strongly implicate the role of IL-2/IL2RA pathway in PSC and provide further confirmation of MST1 association.

Key Words: fine mapping study, genetic association, IL-2/IL-21, IL2RA, primary sclerosing cholangitis

Introduction

Primary sclerosing cholangitis (PSC) is a chronic cholestatic liver disease characterized by progressive inflammation and fibrosis of intra-hepatic and extra-hepatic bile ducts leading to cholestasis and biliary cirrhosis [1]. There is no effective therapy for PSC at present and patients with advanced chronic liver disease ultimately require liver transplantation [2]. PSC has an estimated reported mean incidence of 1.3/100,000 and a point prevalence of 8.5/100,000 [3]. PSC is associated with inflammatory bowel disease (IBD) in 70–80% of cases [4] and up to 25% of patients also suffer with other autoimmune...
diseases [5]. PSC is considered to be a complex disease with circumstantial evidence pointing to a role of an immune-mediated pathology given the high prevalence of other autoimmune diseases and genetic associations within the human leukocyte antigen (HLA) region [6]. It is likely that a complex interaction between environmental and genetic risk factors is required for the development of PSC.

The evidence for a genetic risk in PSC comes from an increased relative sibling risk (\( \lambda_s \)) of developing PSC, which is 9–39 times higher than the general population [7]. In addition, genetic studies have identified association with variants within the HLA region (\( 6p21 \)) which are in the vicinity of class I loci but not confined to this region alone [8,9]. However, genetic associations outside this region are less well defined. Two recent genome-wide association studies (GWAS) identified susceptibility loci at 2q13, 2q35, 3p21, and 13q31 with the most likely candidate genes proposed as BCL2-like 11 (\( BCL2L11 \)), G-protein-coupled bile acid receptor 1 (\( GPBAR1 \)), macrophage stimulating 1 (\( MST1 \)), and glypican 6 (\( GPC6 \)) respectively [8,9]. In addition, several SNPs at interleukin-2 receptor-\( \alpha \) (\( IL2RA \)) locus showed suggestive association in discovery panel of the largest study but failed to replicate, probably due to effect size heterogeneity in the replication panels [9]. Intriguingly, the 4q27 locus harboring the ligand interleukin-2 (\( IL-2 \)) of \( IL2RA \) has also recently been implicated in a candidate gene study in PSC [10].

Genetic risk variants at 4q27 (\( IL-2/IL-21 \)) and 10p15 (\( IL2RA \)) are shared across multiple immune-mediated diseases [11] and are of interest because of the role of \( IL-2 \) signaling in immune tolerance [6]. Furthermore, \( Il-2r^{-/-} \) mice spontaneously develop intestinal and biliary inflammation [12], making \( IL-2/IL2RA \) a highly plausible signaling pathway to be implicated in PSC pathogenesis.

To further characterize and refine the associations at 4q27 (\( IL-2/IL-21 \)) and 10p15 (\( IL2RA \)), a fine mapping study was undertaken in a large UK cohort of PSC cases and healthy UK controls. Single nucleotide polymorphisms (SNPs) at four other non-HLA loci previously reported to be associated with PSC (rs6720394 (2q13), rs12612347 (2q35), rs3197999 (3p21), and rs9524260 (13q31)) were also genotyped to validate their association.

**Subjects and methods**

**Cohort selection**

A total of 1030 PSC cases were recruited from 155 UK NHS hospital trusts including all transplant centers in the UK as part of the UK-PSC consortium (http://psc.medgen.medschl.cam.ac.uk, see online supplementary material for details). Multiregional ethics committee (MREC) approval for the study was granted by the Cambridgeshire 4 National Ethics committee (MREC Number 08/45/008). All PSC cases were diagnosed using standard diagnostic criteria [1]. All cases were of Caucasian ethnicity. The characteristics of the cohort are shown in Table I.

The control population comprised of 5162 individuals from the 1958 British Birth Cohort and National Blood Service samples genotyped as part of the Wellcome Trust Case Control Consortium 2 (WTCCC2) project.

**SNP selection**

Four SNPs previously associated with PSC identified from two GWAS were included. These were rs6720394 (2q13; \( BCL2L11 \)), rs12612347 (2q35; \( GPBAR1 \)), rs3197999 (3p21; \( MST1 \)), and rs9524260 (13q31; \( GPC6 \)) [8,9].

To study the interleukin-2 receptor alpha (\( IL2RA \)) locus, an 80 kbp region was selected on chromosome 10p15 between 6080 and 6160 kbp spanning \( IL2RA \). SNP data were downloaded for the CEU population from HapMap [13] data rel 24/phase II Nov 08, on NCBI B36 assembly, dbSNP b126 and Haplovie v4.2 [14] was used for selecting tag SNPs. The HapMap data were filtered based on the following quality control (QC) thresholds: minimum genotype rate = 80%; maximum Mendelian errors = 1; minimum minor allele frequency (MAF) ≥ 0.05 and Hardy–Weinberg equilibrium (HWE) test \( p \)-value > 0.001. Seventy-seven markers met the QC criteria. The SNP list was refined to 48 tag SNPs using a pairwise tagging approach (\( r^2 \) threshold ≥0.8), force including eight key SNPs which showed suggestive association in the previous GWAS or were associated with other immune-mediated diseases (please refer to online Supplementary Table I for details).

To study the \( IL-2/IL-21 \) locus, a 564 kbp region was selected on chromosome 4q27 between 123236 and 123292 kbp.

**Table I. Patient characteristics.**

| Cohort size | 1030 |
| Ethnic origin | 100% Caucasian (97.5% British) |
| Median age at recruitment | 58 (range, 16–86) years |
| Gender | Males = 650, Females = 380 (M:F = 1.7:1) |
| Inflammatory bowel disease | CrD = 658 (64%); UC = 570 (86%), IC = 5 (1%) |
| Liver transplant recipients | 252 (24%) |

Abbreviations: UC = ulcerative colitis; CrD = Crohn’s disease; IC = indeterminate colitis.
Following imputation, we only retained genotypes with genotype call rate > 95%, according to the MACH output. We excluded two SNPs, rs4833248 (4q27) and rs1996077 (4q27), which showed unacceptably low overall rates of genotype assignment when using this 99% threshold and were discarded from the study. Details of all the SNPs, which passed QC and were analyzed, with method of genotyping (genotyped or imputed) in the control population are given in the online Supplementary Table II.

Quality control

Stringent individual and SNP QC thresholds were applied to both cases and controls.

Cases. Thirty-eight cases were excluded due to low genotyping rate (<90%); the genotyping rate in the remaining cases was > 99%. None of the SNPs showed significant deviation from HWE (p-values > 10^-4). One SNP rs3197999 (3p21; MST1) had a genotype rate of < 95%. This key SNP was found to be associated with PSC in a previous GWAS [9]. For this SNP, genotyping was repeated using TaqMan technology with a genotyping success rate > 98%.

Controls. QC thresholds were applied to 47 genotyped SNPs in 5162 controls before combining the dataset with imputed genotype data for remaining 17 SNPs. Total genotyping rate in controls was > 99% and all the markers had acceptable genotyping rate (> 95%). None of the markers deviated significantly from Hardy–Weinberg equilibrium (p-value > 10^-5). One SNP rs2390352 (4q27) had MAF < 0.05 and was removed from further analysis.

Statistical methods

After QC, 63 SNPs were analyzed in 992 cases and 5162 controls. Allelic association analysis for all the SNPs was performed using logistic regression in PLINK v. 1.07 [16]. We did not use expected allele dosages to allow the analysis of imperfectly imputed control data, as this could result in more precisely measured genotypes in cases (which had definite genotypes) than in controls, thereby, increasing the potential to generate false-positive results. Therefore, we only included control genotypes that had been imputed with > 99% probability. For selected SNPs, previously published summary statistics were used to perform meta-analysis using the Metagen (inverse variance method) package in the R statistical software package (R v 2.13.1). A random effects model was used even though a test for heterogeneity was
not statistically significant for any of the associated SNPs.

We also performed a logistic regression analysis conditioned on the most associated SNPs at 4q27 and 10p15 loci in PLINK v. 1.07, to determine the strength of association of these SNPs in relation to each other.

In addition, we performed sub-group analyses based on autoimmune disease and IBD status (please refer to online Supplementary Material for details).

To correct for multiple testing for the 63 SNPs analyzed, we determined a Bonferroni corrected threshold for a significant p-value at $< 7.9 \times 10^{-4}$ (0.05/63).

## Results

### Allelic association analysis

After quality control, 992 PSC cases were compared with 5162 controls. Allelic association analysis was performed for 63 SNPs using logistic regression in PLINK v1.07. One of the SNPs at four other previously reported non-HLA loci was replicated (3p21, rs3197999; $p = 1.9 \times 10^{-6}$; OR$_{A \text{ vs } G} = 1.28$, 95% confidence interval [CI] (1.16–1.42)) (Table II). In the fine mapping of 10p15, significant associations were observed after correction for multiple testing ($p < 7.9 \times 10^{-4}$), for three SNPs at 10p15 (rs4147359, $p = 2.6 \times 10^{-4}$, OR$_{A \text{ vs } G} = 1.20$, 95% CI (1.09–1.33); rs706778, $p = 4.3 \times 10^{-4}$, OR$_{T \text{ vs } C} = 1.19$, 95% CI (1.08–1.31); and rs7090530, $p = 7.0 \times 10^{-4}$, OR$_{C \text{ vs } A} = 0.84$, 95% CI (0.76–0.93)). A logistic regression analysis conditioned on these three associated SNPs confirmed the association to be dependent on one another (results not shown).

Nominal association ($p < 0.05$) was seen for an additional 10 SNPs at this locus (online Supplementary Table II).

At the 4q27 locus, significant association was observed for rs12511287 ($p = 3.0 \times 10^{-4}$, OR$_A$ vs $T = 1.21$, 95% CI (1.09–1.35)) (Table II). The second most associated SNP at this locus (rs6822844) was associated at nominal significance not robust for multiple corrections in the UK cohort ($p = 1.9 \times 10^{-3}$, OR$_T$ vs $G = 0.81$, 95% CI (0.70–0.92)) (online Supplementary Table II). However, this SNP has recently been shown to be associated with PSC in a cohort of 1,186 northern European PSC patients [10]. Combined analysis using published summary statistics from this study and the UK cohort strengthened the association at rs6822844 ($p = 6.9 \times 10^{-6}$; OR$_T$ vs $G = 0.77$, 95% CI (0.69–0.86)) (online Supplementary Table III). In addition, we performed logistic regression analyses at this locus, conditioned on the two most associated SNPs rs12511287 and rs6822844, which are not in LD with each other (pairwise $r^2 = 0$). Interestingly, both SNPs were independently associated with PSC (see online Supplementary Figure 1). Furthermore, nominal association ($p < 0.05$) was seen for an additional eight SNPs at this locus. A regional association plot for all the SNPs analyzed at 10p15 and 4q27 loci is shown in Figure 1. Association results for all the studied SNPs are given in the online Supplementary Table II.

### Meta-analysis

Summary statistics from the present cohort were combined with results from three other published genetic association studies [8–10] for four previously associated non-HLA loci and selected IL-2/IL2RA SNPs. Results of the combined analysis for selected SNPs reaching significance threshold in the UK cohort are shown in Table II. Genome-wide significance ($p < 5 \times 10^{-8}$) was observed for rs3197999 at 3p21 ($P_{\text{combined}} = 3.8 \times 10^{-12}$,

---

Table II. Results for SNPs associated in the UK cohort and combined analysis.

<table>
<thead>
<tr>
<th>Chr</th>
<th>SNP</th>
<th>Locus</th>
<th>Alleles (minor/major)</th>
<th>MAF (case/control)</th>
<th>p-Value (UK cohort)</th>
<th>OR (95% CI) (UK cohort)</th>
<th>OR (combined)</th>
<th>p-Value (combined)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>rs3197999</td>
<td>MST1</td>
<td>A/G</td>
<td>0.33/0.28</td>
<td>$1.9 \times 10^{-6}$</td>
<td>1.28 (1.16–1.42)</td>
<td>1.38 (1.26–1.51)</td>
<td>3.8 \times 10^{-12}</td>
</tr>
<tr>
<td>4</td>
<td>rs12511287</td>
<td>IL2/IL21</td>
<td>A/T</td>
<td>0.31/0.27</td>
<td>$3.0 \times 10^{-4}$</td>
<td>1.21 (1.09–1.35)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>rs4147359</td>
<td>IL2RA</td>
<td>A/G</td>
<td>0.39/0.34</td>
<td>$2.6 \times 10^{-4}$</td>
<td>1.20 (1.09–1.33)</td>
<td>1.25 (1.16–1.36)</td>
<td>1.5 \times 10^{-8}</td>
</tr>
<tr>
<td>10</td>
<td>rs706778</td>
<td>IL2RA</td>
<td>T/C</td>
<td>0.44/0.40</td>
<td>$4.3 \times 10^{-4}$</td>
<td>1.19 (1.08–1.31)</td>
<td>1.24 (1.14–1.35)</td>
<td>3.4 \times 10^{-7}</td>
</tr>
<tr>
<td>10</td>
<td>rs7090530</td>
<td>IL2RA</td>
<td>C/A</td>
<td>0.36/0.40</td>
<td>$7.0 \times 10^{-4}$</td>
<td>0.84 (0.76–0.93)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Abbreviations: Chr = chromosome, MAF = minor allele frequency, OR = odds ratio, CI = confidence interval. Association results are shown for SNPs associated after correction for multiple testing (Bonferroni corrected threshold for significance $p < 7.9 \times 10^{-4}$). Meta-analysis for selected SNPs was performed by combining summary statistics from a recent PSC GWAS [9], using the Metagen (inverse variance method) package in R statistical software package. A random effects model was used for meta-analysis even though a test for heterogeneity was not statistically significant for any of the associated SNPs. All the reported odds ratios are with reference to minor allele vs. major allele.
OR\textsubscript{A} vs. G = 1.38, 95% CI (1.26–1.51)) and for rs4147359 at 10p15 ($P_{\text{combined}} = 1.5 \times 10^{-8}$, OR\textsubscript{A} vs. G = 1.25, 95% CI (1.16–1.36)). The other previously reported non-HLA SNPs at 2q13, 2q35, and 13q31 were not associated in the UK cohort or combined analysis (results shown in online Supplementary Table III).

**Phenotype-based sub-group analyses**

**Autoimmunity.** Out of 1030 cases, 122 (12%) had one or more co-existing autoimmune diseases and these were removed to perform a sub-group analysis for patients with PSC (with or without IBD) alone ($n = 908$). A list of the autoimmune diseases present in these cases is shown in Table III. The $p$-value results for this sub-group analysis at each SNP are shown in the online Supplementary Table II ($p$-value\textsubscript{(AID excluded)}). Significant association robust for multiple corrections was seen for rs3197999 (3p21; MST1; $p$-value\textsubscript{(AID excluded)} = $1.2 \times 10^{-5}$) and rs12511287 (4q27; IL-2/IL-21; $p$-value\textsubscript{(AID excluded)} = $4.7 \times 10^{-5}$).

**Inflammatory bowel disease.** We used fully quality-controlled dataset ($n = 992$ cases) for this sub-group analysis and identified cases based on their IBD status. Out of 992 cases, 625 (63%) cases had co-existing IBD and 367 suffered only from PSC. In the analysis of PSC cases with co-existing IBD, significant association robust for multiple corrections was seen for rs3197999 (3p21; MST1) ($p$-value\textsubscript{(PSC with IBD)} = $1.2 \times 10^{-5}$). Several other SNPs were associated at nominal significance ($p$-value < 0.05) in both the sub-groups (PSC with IBD; and PSC without IBD), but the associations were not robust for multiple corrections (results shown in the online Supplementary Table II). This was most likely due to the reduced power of this sub-cohort to detect an association given the small sample size in the two groups.

**Discussion**

This study identifies for the first time an association between PSC and SNPs at the 10p15 locus with genome-wide significance and confirms association at 3p21 and 4q27. The most plausible candidate genes at these loci are IL2RA, MST1, and IL-2/IL-21, respectively.

The MST1 SNP, rs3197999, encodes a non-synonymous change p.Arg689Cys in MST1 and is also associated with IBD, suggesting shared genetic risk factors between PSC and IBD [9]. MST1 encodes macrophage-stimulating protein (MSP), which has a critical role in attenuating the inflammatory response by inhibition of lipopolysaccharide-induced inflammatory mediators [18].

At the 10p15 locus, association signal peaks at rs4147359, which is located in the intergenic region between IL2RA and RNA binding motif protein-17

**Table III. List of autoimmune diseases in the PSC cohort.**

<table>
<thead>
<tr>
<th>Autoimmune disease</th>
<th>No. of cases affected (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thyroid disease</td>
<td>79 (7.6%)</td>
</tr>
<tr>
<td>Celiac disease</td>
<td>16 (1.5%)</td>
</tr>
<tr>
<td>Type 1 diabetes (T1D)</td>
<td>13 (1.2%)</td>
</tr>
<tr>
<td>Thyroid disease + T1D</td>
<td>6 (0.5%)</td>
</tr>
<tr>
<td>Thyroid disease + Celiac disease</td>
<td>4 (0.4%)</td>
</tr>
<tr>
<td>Systemic lupus erythematosus</td>
<td>3 (0.3%)</td>
</tr>
<tr>
<td>Sjogrens syndrome</td>
<td>1 (0.1%)</td>
</tr>
</tbody>
</table>
(RBM17). This SNP has been associated previously with type 1 diabetes (T1D) [19]. Other T1D-risk alleles at this locus have been associated with reduced serum concentrations of soluble IL-2RA [19], making IL2RA a plausible candidate gene at this locus, as well as suggesting a possible mechanism for the genetic association. Of note, spontaneous mutation of IL2RA in humans causes systemic autoimmunity and multi-organ inflammation [20].

At the 4q27 locus, the most associated SNP (rs12511287) is located upstream of the IL-21 gene. The second most associated SNP at this locus, rs6822844, located in a noncoding region upstream of IL-2 and downstream of IL-21, has previously been associated with other autoimmune diseases and is considered as a general autoimmune risk locus [21]. We have for the first time shown two independent association signals at this locus marked by rs12511287 and rs6822844 in PSC. A likely candidate gene at 4q27 is IL-2 [10], especially given the strong PSC association with its receptor (IL2RA). IL-2 is a cytokine produced predominantly by activated T cells and exhibits its actions both in an autocrine and paracrine fashion by binding to the IL-2 receptor (IL-2R). It plays a crucial role in modulating the immune response by promoting proliferation, differentiation, and function of activated T cells as well as maintaining the homeostasis and functioning of CD4+CD25+ Foxp3+ T regulatory (Treg) cells [22]. A defect in Treg-cell production is believed to be the main reason for autoimmunity associated with IL-2/IL-2R deficiency as seen in mice models and humans [20,23]. Another gene encoded at 4q27 locus, which also appears to be a plausible candidate gene at this locus, is IL-21. IL-21 plays an important role in immunoglobulin production, has proapoptotic actions on B cells and drives the terminal differentiation of B cells to plasma cells [6]. IL-21 may also have a role in amplifying the gut inflammatory milieu in ulcerative colitis patients, thereby promoting growth of colitis-associated cancer [24]. PSC further increases the risk of colorectal cancer in IBD patients but the mechanism is unknown [25]. It is possible that PSC patients with IBD may have dysregulated IL-21 function, making it an interesting candidate gene at this locus.

Genetic variants at 4q27 (IL-2/IL-21) and 10p15 (IL2RA) are associated with other autoimmune diseases, such as celiac disease, rheumatoid arthritis, type 1 diabetes, multiple sclerosis, and psoriasis [6,11,26]. There is widespread sharing of genetic risk loci (including variants at 4q27 (IL-2/IL-21) and 10p15 (IL2RA)) between these diseases [11], making them non-organ-specific autoimmune disease susceptibility loci. PSC is associated with an increased prevalence of other autoimmune diseases and its association with genetic variants at 4q27 and 10p15 loci could imply not only a shared genetic risk with other autoimmune diseases but also allow us to identify the shared immunological pathways favoring disease development.

The present results refine the known genetic architecture in PSC by confirming MST1, IL2RA and IL-2/IL-21 locus associations, suggesting a role of both innate and adaptive immune responses in PSC pathogenesis. Despite successfully replicating and confirming the associations at these loci, we were not able to refine the association signal to a specific causal variant. It is important to emphasize that genetic association studies (GWAS or candidate gene studies) have a limited utility in identifying true causal genetic variants associated with complex diseases such as PSC. They do allow identification of the most likely biological pathways that may be involved in disease pathogenesis and serve as good starting points to design robust functional studies based on the implicated pathways [27]. PSC is believed to be a multifactorial disease, and genetic polymorphisms not yet discovered are bound to be important along with potential environmental and epigenetic risk factors. Functional studies at each of the associated loci are now required to identify the true causative gene or genes to facilitate rapid translation to the discovery of novel therapeutics.

Acknowledgements

The authors wish to thank all PSC cases and healthy controls for their participation. We acknowledge the Wellcome Trust Case Control Consortium (WTCCC) for providing the healthy control genotype data. We greatly acknowledge all the principal investigators of the UK-PSC consortium and the National Institute for Health Research (NIHR) Comprehensive Clinical Research Network (CCRN) led research nurses for their support received for recruiting the patients enrolled in this study. This study was supported by the UK-based PSC support charity (Registered charity number 1115615) and the Norwegian PSC research center. GFM is a Clinical Research Training Fellow of the Medical Research Council (G0800460). He is also supported by a Raymond and Beverly Sackler Studentship. Study concept and design: B.S., T.H.K., E.M. Case ascertainment and phénotyping: B.S., A.W.M., M.B., The UK-PSC Consortium (see Supplementary material for details), R.W.C., S.M.R. Genotyping of cases: E.E., A.F. Control genotype data ascertainment: G.F.M., R.N.S., H.J.C. Analysis and interpretation of data: B.S., G.F.M., H.J.C., E.M. Manuscript preparation: B.S., R.N.S.,
E.M. Revised the manuscript: S.M.R., H.J.C., E.E., T.H.K., G.J.A. All authors approved of the final version.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References


Supplementary material available online

Supplementary Table I–III
Supplementary Figure 1