Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci

David Ellinghaus1,49, Luke Jostins2, Sarah L Spain2, Adrian Cortes3,4, Jörn Bethune1, Buhm Han5, Yu Rang Park6, Soumya Raychaudhuri7–10, Jennie G Pouget11,12, Matthias Hábenthal1, Trine Folseraas13–16, Yunpeng Wang17, Tomu Õska18–20, Andres Metspalu18, Harm-Jan Westra7–10, Lude Franke21, Tune H Pers7,20,22,23, Rinse K Weersma24, Valerie Collilj24, Mauro D’Amato25,26, Jonas Halfvarson27, Anders Boeck Jensen28, Wolfgang Lieb29,30, Franziska Degenhardt31,32, Andreas J Forstner31,32, Andrea Hofmann31,32, The International IBD Genetics Consortium (IIBDGC)33, International Genetics of Ankylosing Spondylitis Consortium (IGAS)33, International PSC Study Group (IPSCSG)33, Genetic Analysis of Psoriasis Consortium (GAPO)33, Psoriasis Association Genetics Extension (PAGE)33, The International IBD Genetics Consortium (IIBDGC)33, International Genetics of Ankylosing Spondylitis Consortium (IGAS)33, International PSC Study Group (IPSCSG)33, Genetic Analysis of Psoriasis Consortium (GAPO)33, Psoriasis Association Genetics Extension (PAGE)33, The International IBD Genetics Consortium (IIBDGC)33, International Genetics of Ankylosing Spondylitis Consortium (IGAS)33, International PSC Study Group (IPSCSG)33, Genetic Analysis of Psoriasis Consortium (GAPO)33, Psoriasis Association Genetics Extension (PAGE)33

We simultaneously investigated the genetic landscape of ankylosing spondylitis, Crohn’s disease, psoriasis, primary sclerosing cholangitis and ulcerative colitis to investigate pleiotropy and the relationship between these clinically related diseases. Using high-density genotype data from more than 86,000 individuals of European ancestry, we identified 244 independent multidisease signals, including 27 new genome-wide significant susceptibility loci and 3 unreported shared risk loci. Complex pleiotropy was supported when contrasting multidisease signals with expression data sets from human, rat and mouse together with epigenetic and expressed enhancer profiles. The comorbidities among the five immune diseases were best explained by biological pleiotropy rather than heterogeneity (a subgroup of cases genetically identical to those with another disease, possibly owing to diagnostic misclassification, molecular subtypes or excessive comorbidity). In particular, the strong comorbidity between primary sclerosing cholangitis and inflammatory bowel disease is likely the result of a unique disease, which is genetically distinct from classical inflammatory bowel disease phenotypes.

Genome-wide association studies (GWAS) have shown overlap in the genetic susceptibility to human diseases that affect a range of tissues. This overlap is most notable in immune-mediated diseases1,2, including the clinically related conditions ankylosing spondylitis, Crohn’s disease, psoriasis, primary sclerosing cholangitis (PSC) and ulcerative colitis. Comorbidity of these conditions in the same individuals and increased risk of any of these conditions in the family members of cases have been extensively documented3–4. Recently, a large-scale discovery-driven analysis of temporal disease progression patterns using data from an electronic health registry covering the whole population of Denmark found substantial population-wide comorbidity5. This raises the possibility of a hidden molecular taxonomy that differs from the traditional classification of disease by organ or system. Cross-disease genetic studies provide an opportunity to resolve overlapping associations into discrete pathways and explore details of apparently shared etiologies.

In this study, we combined Immunochip genotype data for 52,262 cases and 34,213 controls of European ancestry, currently the largest available genetic data sets in five clinically related seronegative immune-driven phenotypes (ankylosing spondylitis, Crohn’s disease, psoriasis, PSC and ulcerative colitis), to explore the extent of sharing of genetic susceptibility loci. The aims of this cross-phenotype study were (i) to identify subsets of the five phenotypes with shared genetic risk loci using a cross-phenotype meta-analysis approach; (ii) to identify additional susceptibility loci; (iii) to investigate comorbidity and pleiotropy among these phenotypes; and (iv) to improve the understanding of shared pathways and biological mechanisms common to subsets of the phenotypes studied.

RESULTS

Cross-phenotype association analysis

We analyzed Immunochip genotype data for 52,262 cases of ankylosing spondylitis (8,726), Crohn’s disease (19,085), psoriasis (6,530),...
PSC (3,408) and ulcerative colitis (14,413) and 34,213 healthy controls (Supplementary Table 1) using variants with a minor allele frequency (MAF) >0.1% to examine the shared and distinct genetic etiology for these diseases (Online Methods). By using Immunochip-only data, we were able to perform uniform and central quality control of all batches, thus reducing potentially existing batch effects (Online Methods). Next, we used a recently published subset-based meta-analysis approach (SBM)⁶ to exhaustively explore all subsets of disease combinations for the presence of association signals. This method identifies the best subset of non-null studies while in parallel accounting for multiple-hypothesis testing and a fixed control group (Online Methods). By performing primary SBM analyses, we identified 166 genome-wide significant \( P_{\text{SBM}} < 5 \times 10^{-8} \) loci outside the major histocompatibility complex (MHC) region (chromosome 6 region at 25–34 Mb) (Supplementary Fig. 1). Three of these 166 loci \((rs2042011 at MIR1208, rs2812378 at CCL21-FAM205A and rs1893592 at UBASH3A)\) have not been reported previously for any of the five diseases under study and thus are new shared risk loci. SNP associations at UBASH3A \((21q22.3)\) and CCL21 \((9p13.3)\) have been reported previously for other autoimmune disorders⁷. These three new loci would have been missed using single-disease analyses alone. To avoid any loss of power, where variants were only associated with a single phenotype, we carried out single disease versus control subsearches on any SNPs that achieved \( P_{\text{SBM}} < 5 \times 10^{-8} \) in the primary analysis. Using this SBM-directed approach, we identified 27 new genome-wide significant disease associations \( P_{\text{disease}} < 5 \times 10^{-8} \), including 17 new genome-wide significant loci for ankylosing spondylitis, 6 loci for Crohn's disease and 4 loci for PSC (Fig. 1, Supplementary Fig. 2 and Supplementary Table 2). Twenty-four of these 27 associations were also genome-wide significant in the primary SBM analyses \( P_{\text{SBM}} < 5 \times 10^{-8} \), thus leading to a total of 169 non-MHC risk loci. To identify additional independent association signals within the 169 non-MHC risk loci, we performed a stepwise conditional SBM analysis following a recently published stepwise conditional SBM fine-mapping approach⁹ (Online Methods). In total, we identified 244 independent association signals, with 187 signals shared by at least two of the five diseases under study (Supplementary Figs. 3–5 and Supplementary Table 3). We estimated the heritability explained by these 244 variants for each disease (Fig. 2) and for all pairwise disease comparisons (Supplementary Fig. 6). The ten pairwise comparisons of disease-associated alleles showed diverse patterns of sharing with respect to the size and direction of allelic effects and the number of unique associations (Supplementary Fig. 6).

**Functional annotation of associated variants**

We functionally annotated the 244 risk alleles from the 169 distinct loci (Online Methods). For 210 association signals (86.1%), the lead variant was within 10 kb of a known gene, and 34 signals were classified as occurring in intergenic regions (>10 kb distant to a gene) (Supplementary Table 4). The analysis identified 16 coding variants (14 missense, 1 frameshift and 1 splice donor) in genes that were previously implicated in immune-mediated diseases (Supplementary Table 5). Eight of these variants (located in PTPN22, GPR35, MST1, CD6, NOD2 (two variants), TYK2 and CARD9) have been associated before with one of the traits included in this study, and six \((GCKR, IFIH1 (two variants), SH2B3, SMAD3 and TYK2)\) have been associated with another phenotype, either listed in the GWAS catalog¹⁰ or in ImmunoBase. Two of the genes carrying a coding variant \((TLR4 and PRKCC)\) have previously been suggested as candidate loci, but robust association signals were still lacking (Supplementary Table 5).

We further checked for variants in high linkage disequilibrium \((LD; r^2 > 0.8)\) with the identified variants using 1000 Genomes Project haplotypes and found that, in total, 46 of the identified signals were highly correlated with 57 coding variants (48 missense, 2 stop-gain, 3 splice-region, 1 frameshift and 3 regulatory variants; Supplementary Table 5). We found that 40 of the 57 coding variants, from 30 loci, had been described in previous GWAS or Immunochip studies involving one of the traits included in this study or another phenotype. Additionally, a further nine variants have been mentioned as candidate variants in autoimmune disease publications. Eight coding variants (seven missense and one stop gain; all in high LD with our lead variants), located in ENFA1, FCGR2A, HSPA6, C7orf72 and FAM118A, have not been described before in relation to any immune-mediated phenotype.

**eQTL analysis in peripheral blood**

Analyses of cis expression quantitative trait locus (eQTL) microarray data from whole peripheral blood samples of an independent control cohort comprising 2,360 unrelated individuals¹¹,¹² (Online Methods) identified cis effects for 132 (false discovery rate (FDR)-adjusted \( P_{\text{FDR}} < 0.05; \text{Supplementary Table 6})\) of the 244 disease-associated SNPs from Supplementary Table 3a. Five of these represent the best eQTL SNP, and another five represent the best secondary eQTL SNP independent from the best eQTL SNP at a given locus.

**Pathway, cell type and annotation enrichment analyses**

We tested for enrichment between SNPs in associated loci and various types of genomic annotations using GoShifter¹³. We used 620 different annotations from the US National Institutes of Health (NIH) Roadmap Epigenomics¹⁴ and FANTOM5 (ref. 15) projects to look for enrichment of histone modifications and expressed enhancers, respectively (Supplementary Tables 7 and 8). The results from the SBM association analysis were separated into groups including all 244 identified variants, variants shared by three or more phenotypes and variants associated with a single phenotype (Supplementary Table 9). In the Roadmap Epigenomics enrichment analysis, using a significance threshold of \( P < 1 \times 10^{-3}\), the inflammatory bowel disease (IBD) and psoriasis phenotype subsets showed enrichment for acetylation of histone H3 at lysine 27 \((H3K27ac)\) in CD3 primary cells and for H3K27ac in adipose tissue, respectively (Supplementary Table 10). The ‘all variants’ \((n = 244)\) analysis showed enrichment for trimethylation of histone H3 at lysine 4 \((H3K4me3)\); for which the largest number of cell types were analyzed by the Roadmap Epigenomics consortium and which highlights transcribed promoters and transcription start sites (TSSs)¹⁶ in the HUES64 cell line (representing undifferenntiated cells), CD34⁺ cells (bone marrow cells) and natural killer cells (CD56). The FANTOM5 data analysis showed enrichment for enhancers expressed in T cells (Crohn's disease and the ‘all variants’ group) and also natural killer cells for Crohn's disease. However, only the enrichment for enhancers expressed in T cells in the ‘all variants’ group met the significance threshold of \( P = 0.05/620 = 8.06 \times 10^{-4}\) required with Bonferroni correction.

To test which candidate genes from the associated loci (Supplementary Table 3a) are highly expressed in which tissues and to define disease relationships at the expression level, we conducted pathway and tissue/cell type enrichment analyses using DEPICT¹⁷, with 77,840 microarray expression profiles from human, rat and mouse and 209 tissue/cell type annotations¹⁸ (Online Methods). Even when correcting for the biased Immunochip gene content, our DEPICT results confirmed that the genes from the 169 non-MHC genome-wide significant susceptibility loci reported herein show greatest relevance for the regulation of immune response pathways (Supplementary Fig. 7) and the hematopoietic system (Supplementary Fig. 8).
We further generated a protein-protein interaction (PPI) network (111 gene nodes and 65 edges; Supplementary Fig. 9) based on five prioritized gene sets for the ankylosing spondylitis, Crohn’s disease, psoriasis, PSC and ulcerative colitis SNP sets (Supplementary Table 9), from DEPICT analyses and reference PPI data from ConsensusPathDB [19] (Online Methods). We observed that 36 gene nodes from this PPI network were connected in one single large component (Supplementary Fig. 9). We then evaluated the potential role of these genes for their ‘druggability’ by linking genes within this core network to drugs using DrugBank (Online Methods). Because the nature and effect of the interaction between a drug and an encoded protein is mostly unknown (for example, some drugs we identified have effects opposite to what we aim for), we performed a manual literature search to assess which of the identified drugs show evidence or could potentially be promising for any of the diseases under study by using PubMed (last search 1 July 2015) and ClinicalTrials.gov. All drugs were selected on the basis of evidence from phase I, II and/or III randomized clinical trials (RCTs) or published animal studies. Nine drug target genes overlapped with the 36 genes from the core network (Fig. 3). Although further investigations are necessary, we propose that target gene–drug combinations selected by this approach could represent promising candidates for novel drug discovery in the treatment of ankylosing spondylitis, Crohn’s disease, psoriasis, PSC and ulcerative colitis. For example, novel CCR2 antagonists such as MLN-1202 and the CCR5 antagonists INCB9471 and AMD-070 are potential new drugs for treatment of ankylosing spondylitis, Crohn’s disease, psoriasis, PSC and ulcerative colitis.

Bayesian multinomial regression for model selection
To compare different disease models for each of the 244 risk variants while accounting for the different sample sizes for the diseases, we used Bayesian multinomial regression. The aim is to estimate the posterior probability ($P_{\text{model}}$) for each disease model conditional on the genotype and phenotype data that were observed (Online Methods). A disease model is a set of diseases with which a given locus is associated (has a nonzero log-transformed odds ratio; for example, ‘associated with Crohn’s disease and ulcerative colitis but not with ankylosing spondylitis, psoriasis or PSC’ is one disease model). There are a total of 32 possible disease models for the five phenotypes, which include the null model (‘not associated with any disease’). The Bayesian setting naturally handles the different uncertainties on the effect sizes for each disease resulting from their different sample sizes and powers.

We found 66 signals (59 non-MHC loci) with a best $P_{\text{model}} \geq 60\%$, including 14 loci (with closest genes $SH2B3$, $UBE2L3$, $TNP2$, $IL2RA$, $DNMT3B$, $CXC2R$, $CDKAL1$, $CARD9$, $MST1$, $ZMIZ1$ and $ETS1$) with $P_{\text{model}} \geq 0.8$ (Supplementary Table 3b) when assuming that each sharing model is given the same probability (uniform prior across all models; Online Methods). However, because previous studies suggested that the structure of sharing in susceptibility is non-uniform [2], we calculated posteriors for each model for each risk variant under six different priors and took a vote of the highest prior across all models; Online Methods). However, because previous studies suggested that the structure of sharing in susceptibility is non-uniform [2], we calculated posteriors for each model for each risk variant under six different priors and took a vote of the highest prior across all models; Online Methods). We then counted how many priors voted for that model and calculated the posterior probability ($P_{\text{model}}$) for each disease model conditional on the genotype and phenotype data that were observed (Online Methods). A disease model is a set of diseases with which a given locus is associated (has a nonzero log-transformed odds ratio; for example, ‘associated with Crohn’s disease and ulcerative colitis but not with ankylosing spondylitis, psoriasis or PSC’ is one disease model). There are a total of 32 possible disease models for the five phenotypes, which include the null model (‘not associated with any disease’). The Bayesian setting naturally handles the different uncertainties on the effect sizes for each disease resulting from their different sample sizes and powers.

We found 66 signals (59 non-MHC loci) with a best $P_{\text{model}} \geq 60\%$, including 14 loci (with closest genes $SH2B3$, $UBE2L3$, $TNP2$, $IL2RA$, $DNMT3B$, $CXC2R$, $CDKAL1$, $CARD9$, $MST1$, $ZMIZ1$ and $ETS1$) with $P_{\text{model}} \geq 0.8$ (Supplementary Table 3b) when assuming that each sharing model is given the same probability (uniform prior across all models; Online Methods). However, because previous studies suggested that the structure of sharing in susceptibility is non-uniform [2], we calculated posteriors for each model for each risk variant under six different priors and took a vote of the highest prior across all models; Online Methods). We then counted how many priors voted for that model and calculated the posterior probability ($P_{\text{model}}$) for each disease model conditional on the genotype and phenotype data that were observed (Online Methods). A disease model is a set of diseases with which a given locus is associated (has a nonzero log-transformed odds ratio; for example, ‘associated with Crohn’s disease and ulcerative colitis but not with ankylosing spondylitis, psoriasis or PSC’ is one disease model). There are a total of 32 possible disease models for the five phenotypes, which include the null model (‘not associated with any disease’). The Bayesian setting naturally handles the different uncertainties on the effect sizes for each disease resulting from their different sample sizes and powers.

We found 66 signals (59 non-MHC loci) with a best $P_{\text{model}} \geq 60\%$, including 14 loci (with closest genes $SH2B3$, $UBE2L3$, $TNP2$, $IL2RA$, $DNMT3B$, $CXC2R$, $CDKAL1$, $CARD9$, $MST1$, $ZMIZ1$ and $ETS1$) with $P_{\text{model}} \geq 0.8$ (Supplementary Table 3b) when assuming that each sharing model is given the same probability (uniform prior across all models; Online Methods). However, because previous studies suggested that the structure of sharing in susceptibility is non-uniform [2], we calculated posteriors for each model for each risk variant under six different priors and took a vote of the highest prior across all models; Online Methods). We then counted how many priors voted for that model and calculated the posterior probability ($P_{\text{model}}$) for each disease model conditional on the genotype and phenotype data that were observed (Online Methods). A disease model is a set of diseases with which a given locus is associated (has a nonzero log-transformed odds ratio; for example, ‘associated with Crohn’s disease and ulcerative colitis but not with ankylosing spondylitis, psoriasis or PSC’ is one disease model). There are a total of 32 possible disease models for the five phenotypes, which include the null model (‘not associated with any disease’). The Bayesian setting naturally handles the different uncertainties on the effect sizes for each disease resulting from their different sample sizes and powers.

We found 66 signals (59 non-MHC loci) with a best $P_{\text{model}} \geq 60\%$, including 14 loci (with closest genes $SH2B3$, $UBE2L3$, $TNP2$, $IL2RA$, $DNMT3B$, $CXC2R$, $CDKAL1$, $CARD9$, $MST1$, $ZMIZ1$ and $ETS1$) with $P_{\text{model}} \geq 0.8$ (Supplementary Table 3b) when assuming that each sharing model is given the same probability (uniform prior across all models; Online Methods). However, because previous studies suggested that the structure of sharing in susceptibility is non-uniform [2], we calculated posteriors for each model for each risk variant under six different priors and took a vote of the highest prior across all models; Online Methods). We then counted how many priors voted for that model and calculated the posterior probability ($P_{\text{model}}$) for each disease model conditional on the genotype and phenotype data that were observed (Online Methods). A disease model is a set of diseases with which a given locus is associated (has a nonzero log-transformed odds ratio; for example, ‘associated with Crohn’s disease and ulcerative colitis but not with ankylosing spondylitis, psoriasis or PSC’ is one disease model). There are a total of 32 possible disease models for the five phenotypes, which include the null model (‘not associated with any disease’). The Bayesian setting naturally handles the different uncertainties on the effect sizes for each disease resulting from their different sample sizes and powers.
risk variant (Supplementary Table 3c). On the basis of this consensus-finding process of merging results from six different priors, we identified 34 signals (31 non-MHC loci) with a best MeanProb model ≥60%, including 12 loci (with closest genes SH2B3, IL2RA, IFIH1, NFKB1 and TYK2) with MeanProbmodel ≥0.8, suggesting that we correctly identified the disease model (Table 1). Of the 34 associations with MeanProbmodel ≥0.6, 25 signals had five diseases involved, 6 signals had four diseases involved and 3 signals were unique to a single disease. Some of these disease sets showed different directions of effect (risk versus protective), heterogeneity of odds ratios (P < 0.01), or both, for the diseases involved (Table 1 and Supplementary Fig. 6).

**Distinguishing pleiotropy from heterogeneity**

Statistically significant temporal comorbidity (disease A followed by disease B within a 5-year time frame of disease A, or vice versa) among the five diseases studied was confirmed for eight of ten possible pairs of diseases (P < 0.05/823,606 = 1.21 × 10⁻⁴) after screening 823,606 directed pairs of diagnoses from an electronic health registry covering the whole population of Denmark. Consistent with previous reports, we further observed high comorbidity rates among our patients (Supplementary Table 12); that is, patients had more than one disease at the time of last diagnosis. This may occur because of pleiotropy (sharing of risk alleles by disease A and disease B) or heterogeneity (a subgroup of disease A cases has a higher load of risk alleles for disease B). Heterogeneity can occur as the result of many different scenarios, including diagnostic misclassifications, molecular subtypes and excessive comorbidity. We evaluated whether pleiotropy or heterogeneity best explained the high comorbidity rates among the five diseases studied using BUHMBOX (Online Methods). BUHMBOX detects heterogeneity by calculating the cross-locus correlation of disease B–associated loci among disease A cases; a nonzero correlation is indicative of heterogeneity. We calculated the statistical power of BUHMBOX to detect various proportions of sample heterogeneity for all disease pairs (Online Methods). For 18 of the 20 pairs of diseases, we had >50% power to detect 20% sample heterogeneity (Supplementary Fig. 10 and Supplementary Table 13). Because BUHMBOX has high power for these pairs,
non-significant BUHMBOX results would strongly suggest that the genetic risk score (GRS) association is likely due to pleiotropy rather than heterogeneity.

First, to quantify genetic sharing for each of the 20 possible pairs of five diseases, we used a GRS approach (Online Methods). We calculated GRSs for disease B (using known risk alleles, weighted by effect size) for all individuals in the disease A sample and tested the association of the GRSs with disease A status using logistic regression. The GRSs test for enrichment of disease B alleles in disease A cases and are expected to be significant in the presence of both pleiotropy and heterogeneity. As expected, we observed highly significant associations between disease B GRSs and disease A status for almost every possible pair (Supplementary Table 14), which demonstrates strong sharing of risk alleles by the different immune-mediated diseases.

We then tested whether this observed genetic sharing was due to true pleiotropy or heterogeneity using BUHMBOX\textsuperscript{20}. In the setting of pleiotropy, pleiotropic disease B risk alleles are shared across all disease A cases, whereas in heterogeneity only a subset of disease A cases share disease B risk alleles. This leads to cross-locus correlations between disease B–associated loci being positive in the presence of heterogeneity but not in the case of pleiotropy. BUHMBOX calculates the cross-locus correlation between disease B–associated loci in disease A cases and determines whether this correlation is significantly nonzero. We calculated cross-locus correlations for all 20 disease pairs (Online Methods). We did not observe significant interlocus correlations (Supplementary Table 14), despite having high statistical power for many pairs (Supplementary Fig. 10 and Supplementary Table 13). Our findings suggest that the overall GRS association between the five immune diseases investigated is likely due to pleiotropy.
Table 1 Bayesian logistic regression analysis identified 31 loci with 34 independent associations for which we determined a specific disease model constellation with high certainty (MeanProbmodel ≥0.6)

| Locus| Signal| Chr. | Locus position| SNP | Nearby gene| OR (AS)| OR (CD)| OR (PS)| OR (UC)| Best model | Vote count | Vote % |
|------|-------|-----|---------------|-----|------------|--------|--------|--------|--------|------------|-----------|-------|
| 217  | 1     | 48  | 45,843,311    | rs314504 | HBB     | 0.95   | 0.77   | 0.63   | 1.10   | PS         | 10        | 0.92  |
| 131  | 1     | 11  | 11,556,972    | rs314504 | HBB     | 0.96   | 0.77   | 0.63   | 1.10   | PS         | 10        | 0.92  |
| 152  | 1     | 11  | 11,646,972    | rs314504 | HBB     | 0.96   | 0.77   | 0.63   | 1.10   | PS         | 10        | 0.92  |
| 217  | 1     | 48  | 45,843,311    | rs314504 | HBB     | 0.96   | 0.77   | 0.63   | 1.10   | PS         | 10        | 0.92  |

Immunochip-wide coheritability analysis

To estimate Immunochip-wide pleiotropy (the genetic variation and covariation between pairs of diseases in liability that is tagged by SNPs represented on the Immunochip), we applied univariate and bivariate linear mixed-model heritability methods (Online Methods). The relationships between disorders are expressed as SNP-based coheritability estimates (Table 1). For correlation values including MHC variants, see Supplementary Table 15. No coheritability was observed with the non-immune-mediated diseases studied here.

DISCUSSION

By combined assessments of Immunochip genotyping data sets from 52,262 patients with five closely associated conditions (ankylosing spondylitis, Crohn's disease, psoriasis, PSC and ulcerative colitis; all seronegative inflammatory diseases according to clinical definition) and 34,213 healthy controls, we were able to delineate the genetic overlap between the conditions. A key outcome of the overlap analysis is that, despite the profound pleiotropy, clear demarcations of the genetic risk for the individual conditions exist. Implicit to this, and...
The genetic landscape of ankylosing spondylitis (AS) is complex and multifactorial. While previous studies have identified several genetic variants associated with AS, such as the intronic SNP rs1801274 in FCGR2A (odds ratio (OR) = 1.09), these findings have been inconsistent and the genetic architecture remains largely unexplored.

Most evidence suggests that AS is caused by a combination of genetic and environmental factors. Previous genome-wide association studies (GWAS) have identified more than 30 loci associated with AS, with the strongest signals coming from the major histocompatibility complex (MHC) region on chromosome 6. However, the majority of these loci explain only a small proportion of the heritability of AS, suggesting that additional genetic factors remain to be discovered.

Despite the genetic complexity of AS, there is evidence of pleiotropy, with AS being associated with other diseases such as Crohn's disease (CD) and ulcerative colitis (UC). This suggests that some genetic factors may contribute to both AS and these gastrointestinal disorders.

The role of genetic variants in AS is likely mediated through inflammation and immune response pathways. For example, the genetic variant rs1801274 in FCGR2A, which encodes the Fc receptor for IgG2, is implicated in the regulation of immune responses. Variants in other genes, such as TNF, IL10, and TYK2, have also been associated with AS and other inflammatory diseases.

In summary, the genetic landscape of AS is characterized by the involvement of multiple genes and pathways, and the interaction between genetic and environmental factors. Further research is needed to identify additional genetic variants and to understand the mechanisms underlying the complex genetic architecture of AS.
with the hypothesis that PSC-IBD is a unique disease that shares some genetic factors with ulcerative colitis but is distinct from classical IBD phenotypes. This hypothesis is further supported by the observation that PSC-IBD shows clinical differences from classical IBD and requires specialized management; in comparison to IBD, PSC-IBD has a higher rate of pancolitis with ileitis and rectal sparing, as well as a higher incidence of colorectal cancer.

Our results from testing of enrichment between multidisease signals and large-scale expression data sets together with epigenetic and expressed enhancer profiles further reflect this extensive pleiotropy and mainly highlight perturbations in immune response pathways and blood cell tissues. However, we could not pinpoint which genomic features and which cells a variant influences. We propose that larger gene expression data sets for the disease-relevant tissues and cell types from affected individuals should be generated to allow for high resolution and more eQTL studies, as eQTLs are often cell type specific. Further, the discovery of multiple additional genetic associations increases the power of such analyses to define pathways and cell types involved in specific diseases.

In summary, we performed the largest systematic cross-disease genetic study for chronic immune-mediated diseases thus far. Using novel cross-phenotype analytical methodologies, we identified 17 new genome-wide significant susceptibility loci for ankylosing spondylitis, 6 loci for Crohn's disease and 4 loci for PSC, and we identified 3 new, as yet unreported risk loci for the diseases under study. With this, the numbers of known ankylosing spondylitis, IBD and PSC risk loci are increased to 48, 206 and 20, respectively. Because of lower coverage at unselected regions on the Immunochip, imputed GWAS data would further increase statistical power to identify new shared associations outside established risk loci in future studies. Future cross-disease studies of a wider range of phenotypes, in combination with more sophisticated fine-mapping studies on individual diseases and specific layers of multi-omics data sets, are needed to provide another layer of information for a potential new disease classification based on molecular genetic profiles. Whereas most cross-disease studies employ patient panels that were manually curated for single phenotypes and often rely on questionnaire data, future studies could employ even larger collections of hospitalized patients, for whom exhaustive electronic medical patient records and array data exist. Moreover, longitudinal data from electronic health charts could pinpoint further comorbidities that should be included in a more systematic next-generation cross-disease approach.

Accession codes. Data access was granted by the management committees of the main disease consortia (the International IBD Genetics Consortium (IBDGC), the International Genetics of Ankylosing Spondylitis Consortium (IGAS), the International PSC Study Group (IPSSCG), the Genetic Analysis of Psoriasis Consortium (GAPC) and the Psoriasis Association Genetics Extension (PAGE)). The genotype data are not freely accessible, but access can be obtained by submitting an application to the respective management committees, institutions or data owners.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS
We thank all patients with ankylosing spondylitis, Crohn's disease, PSC, psoriasis and ulcerative colitis, their families, healthy control individuals and clinicians for their participation in this study. We thank T. Wesse, T. Henke, S. Sedghpour Sabet, R. Vogler, G. Jacobs, I. Urbach, W. Albrecht, V. Pelkonen, K. Holm, H. Dahlen Sollid, B. Woldseth, J.A. Anmarkrud and L. Wenche Torjesjensen for expert help. F. Braun, W. Kreisel, T. Berg and R. Günther are acknowledged for contributing German patients with PSC. B.A. Lie and the Norwegian Bone Marrow Donor Registry at Oslo University Hospital, Rikshospitalet, in Oslo and the Nord-Trøndelag Health Study (HUNT) are acknowledged for sharing the healthy Norwegian controls.

This work was supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the e:med research and funding concept (SysImmune grant 01ZX1006A). This project received infrastructure support from Deutsche Forschungsgemeinschaft (DFG) Excellence Cluster 306 ‘Inflammation at Interfaces’ and the PopGen Biobank. A.F. receives an endowment professorship by the Foundation for Experimental Medicine (Zurich, Switzerland). The Estonian Genome Center at the University of Tartu (EGCUT) received targeted financing from Estonian Research Council grant IUT20-60, the Center of Excellence in Genomics (EXCEGEN) and the University of Tartu (SPIGVAREN). We acknowledge EGCUT technical personnel, especially V. Soo and S. Smit. Data analyses were carried out in part at the High-Performance Computing Center of the University of Tartu. We acknowledge support from the UK Department of Health via National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre awards to Guy's and St Thomas' National Health Service (NHS) Foundation Trust in partnership with King's College London and to Addenbrooke's Hospital in partnership with the University of Cambridge. The study was supported by the German Federal Ministry of Education and Research (BMBF), within the context of National Genome Research Network 2 (NGFN-2), National Genome Research Network plus (NGFNplus) and the Integrated Genome Research Network (IG) Moods (grants 01GS08144 and 01GS08147). R.K.W. is supported by a VIDI grant (016.136.308) from the Netherlands Organization for Scientific Research (NWO). J.H. was supported by the Swedish Research Council (521-2011-2764). This work is supported in part by funding from the US NIH (1R01AR063759 (S.R.), 5U01GM092691-05 (S.R.), 1U12AR067677-01 (S.R.), U19AI112241-01 (S.R.) and 1R01DK084960-05 (K.N.L)) and Doris Duke Charitable Foundation grant 2013097. A.B.J. and S.B. acknowledge funding from the Novo Nordisk Foundation (grant NN14CC0001) and the H2020 project MedBioinformatics (grant 633414). The study was supported by the Norwegian PSC Research Center. We thank G. Trynka for assistance in setting up GoShifter.

AUTHOR CONTRIBUTIONS
D.E., L.I., S.L.S., A.C., J.B., B.H., Y.R.P., J.G.P., S.R., Y.W., T.E., H.-J.W., L.F., T.H.P., R.K.W., V.C., O.A.A., A.B.J., S.B. and A.M.D. performed statistical and computational analyses. M.H. performed computational analyses. T.F., A.M., T.H.P., R.K.W., V.C., O.A.A., A.B.J., S.B. and A.M.D. performed statistical and computational analyses. M.H. performed computational analyses. T.F., A.M., M.D.A., J.H., W.L., F.D., A.I.E., A.H., S.S., U.M., B.D.J., K.N.L., R.C.T., S.W., M.W., E.E., I.T.E., J.W.N.R, and M.A.R. were involved in study subject recruitment and assembling phenotypic data. D.E. wrote the draft of the manuscript. D.E., D.P.M., T.H.K., J.C.B., M.P., M.A.B. and A.F. conceived, designed and managed the study. All authors reviewed, edited and approved the final manuscript.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

METHODS
Methods and any associated references are available in the online version of the paper.

1. Jostins, L. et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 491, 119–124 (2012).
2. Parkes, M., Cortes, A., van Heel, D.A. & Brown, M.A. Genetic insights into common pathways and complex relationships among immune-mediated diseases. Nat. Rev. Genet. 14, 661–673 (2013).
3. Najarian, D.J. & Gottlieb, A.B. Connections between psoriasis and Crohn’s disease. J. Am. Acad. Dermatol. 58, 805–823 (2008).

4. Loftus, E.V. Jr. et al. PSC-IBD: a unique form of inflammatory bowel disease associated with primary sclerosing cholangitis. Gut 54, 91–96 (2005).

5. Jensen, A.B. et al. Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. Nat. Commun. 5, 4022 (2014).

6. Bhattacharjee, S. et al. A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. Am. J. Hum. Genet. 90, 821–835 (2012).

7. Stahl, E.A. et al. Global association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat. Genet. 42, 508–514 (2010).

8. Grant, S.F. et al. Follow-up analysis of genome-wide association data identifies novel loci for type 1 diabetes. Diabetes 58, 290–299 (2009).

9. Wang, Z. et al. Impression and subassociation analysis across different cancer types identifies multiple independent risk loci in the TERT-CLPTM1L region on chromosome 5p15.33. Hum. Mol. Genet. 23, 6611–6633 (2014).

10. Hindorff, L.A. et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc. Natl. Acad. Sci. USA 106, 9362–9367 (2009).

11. Fehrmann, R.S. et al. Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. PLoS Genet. 7, e1002197 (2011).

12. Nia, M. et al. Genetic structure of Europeans: a view from the North-East. PLoS One 4, e5472 (2009).

13. Trynka, G., Bjerke, M., Bengtsson, C. & Finnila, M. Typing of diseasing the effects of colocalizing genomic annotations to functionally prioritize non-coding variants within complex-trait loci. Am. J. Hum. Genet. 97, 139–152 (2015).

14. Roadmap Epigenomics Consortium. Integrative analysis of 111 reference human epigenomes. Nature 518, 317–330 (2015).

15. FANTOM Consortium and the RIKEN PMI and CLST (DGt). A promoter-level mammalian expression atlas. Nature 507, 462–470 (2014).

16. Ernst, J. et al. Mapping and analysis of chromatin state dynamics in nine human cell types. Nature 473, 43–49 (2011).

17. Pers, T.H. et al. Biological interpretation of genome-wide association studies using predicted gene functions. Nat. Commun. 6, 5890 (2015).

18. Fehrmann, R.S. et al. Gene expression analysis identifies global gene dosage sensitivity in cancer. Nat. Genet. 47, 115–125 (2015).

19. Kamburov, A. et al. ConsensusPathDB: toward a more complete picture of cell biology. Nucleic Acids Res. 39, D712–D717 (2011).

20. Han, B. et al. Using genotypic data to distinguish pleiotropy from heterogeneity: deciphering coheritability in autoimmune and neuropsychiatric diseases. bioRxiv doi:10.1101/030783 (6 November 2015).

21. Lee, S.H., Way, N.R., Goddard, M.E. & Visscher, P.M. Estimating missing heritability in case-control studies. Am. J. Hum. Genet. 88, 294–305 (2011).
Online Methods

Study Subjects. All DNA samples included in the study (Supplementary Table 1) were genotyped using the Illumina Immunochip custom genotyping array, a targeted high-density genotyping array with comprehensive coverage of 10,000 Genomes Project SNPs within 186 autoimmune disease-associated loci. Crohn's disease and ulcerative colitis case and control cohorts were collected from 15 countries across Europe, North America and Australia and have previously been described. Initially, 19,761 Crohn's disease cases, 14,833 ulcerative colitis cases and 28,999 controls of European ancestries from IBDGC were included in the study. Genotyping of the IBDGC cohorts was performed in 11 different batches (34 batches before quality control) across 11 different genotyping centers. The initial ankylosing spondylitis case-control collection (two main batches) consisted of 10,417 cases and 12,338 controls of European ancestry and was described previously. All ankylosing spondylitis case-genotyping was performed at one center (University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Australia). 6,577 psoriasis cases and 15,085 control samples (two main batches) were collected from 13 countries across Europe and North America. Recruitment of 3,789 PSC cases and 25,079 controls (two main batches) was performed in 14 countries in Europe and North America. Because most control samples were shared by different disease consortia, we identified the set of non-overlapping (unique) control samples (Supplementary Table 1). In addition, 2,019 schizophrenia cases, 1,140 bipolar disorder cases and 589 major depressive disorder cases were collected from different centers in Germany in the context of the MoodDS consortium. These samples have been genotyped at the Life and Brain Centre (Bonn, Germany).

Written, informed consent was obtained from all study participants, and the institutional ethical review committees of the participating centers approved all protocols.

Immunochip genotype calling and quality control. Initial genotype calling was performed with the Illumina GenomeStudio GenTrain 2.0 software and the custom-generated cluster file of Trynka et al. (based on an initial clustering of 2,000 UK samples and subsequent manual readjustment of cluster positions). On the basis of normalized intensity information, we removed samples detected as intensity outliers (>4 s.d.). On the basis of initial genotype data, we further removed samples with call rate <90% using PLINK. To identify ancestry outliers (subjects of non-European ancestry), we performed a principal-component analysis (PCA) with EIGENSTRAT and a set of 210 HapMap founder samples and projected Immunochip samples on the principal-components axes using a set of 14,484 independent (MAF >0.05) SNPs, excluding X and Y chromosomes, SNPs in LD (leaving no pairs with r2 >0.2) and 11 high-LD regions as described by Price et al. OptiCall genotype calling was performed with a Hardy-Weinberg equilibrium P-value threshold of 1 × 10^{-5} for each batch, Hardy-Weinberg equilibrium blanking disabled and a genotype call threshold of 0.7. Hardy-Weinberg equilibrium was calculated with conditioning on predicted (European) ancestry, and related individuals were removed from this calculation.

After genotype calling, a unified quality control procedure was conducted across 40 genotyping batches. We tested for significantly different allele frequencies of variants across the batches from a particular disease or the control group (with at most one batch being removed) with an FDR threshold of 0.01 (Supplementary Fig. 14). Variants that had >2% missing data, had a MAF <0.1% in either of the different disease sets or in controls, had different missing genotype rates in affected and unaffected individuals (P_{Fisher} < 1 × 10^{-5}) or deviated from Hardy-Weinberg equilibrium (with an FDR threshold of 1 × 10^{-5} in controls) across the entire collection with at most one batch being removed (Supplementary Fig. 15a) or (ii) falling below in two single batches (Supplementary Fig. 15b)) were excluded. Samples that had >2% missing data and overall increased or decreased heterozygosity rates were removed (Supplementary Fig. 16). For robust duplicate/relatedness testing (identity-by-state (IBS) and identity-by-descent (IBD) estimation) and population structure analysis, we used a pruned subset of 14,484 independent SNPs. Pairwise percentage IBD values were computed using PLINK. By definition, Z0: P(IBD = 0), Z1: P(IBD = 1), Z2: P(IBD = 2), Z0 + Z1 + Z2 = 1 and PL_HAT: P(IBD = 2) + 0.5 × P(IBD = 1) (proportion IBD). One individual (the one showing greater missingness) from each pair with PL_HAT >0.1875 was removed.

To resolve within-Europe relationships and to test for population stratification, the remaining quality-filtered 52,262 cases and 34,213 unique controls were tested using the PCA method, as implemented in FlashPCA. PCA identified no outliers with non-European ancestry (Supplementary Fig. 17a-c). We computed Tracy–Widom statistics to evaluate the statistical significance of each principal component identified by PCA and identified the top seven axes of variation significant at P_{TW} < 0.05 (Supplementary Table 16). 130,052 quality-filtered polymorphic variants with MAF >0.1% and 52,262 cases and 34,213 controls were available for analysis.

Cross-phenotype association analysis. We conducted primary association analysis based on subsets (ASSET) methodology. Even after adjusting for the large number of comparisons, the subset-based meta-analysis (SBM) method maintains similar type I error rates as for standard meta-analysis. This method offers a substantial power increase (sometimes approaching between 100–500%) as compared to standard univariate meta-analysis approaches, where the (heterogeneous) effect of a specific SNP is not exclusively restricted to a single disease. Under the assumption that association signals from shared risk loci based on positional overlap are tagging the same causal variant for different diseases, the (unconditioned) SBM approach improves power as compared to standard fixed-effects meta-analysis methodology. For the situation that distinct variants within shared susceptibility regions may confer independent effects for individual diseases, the conditional SBM approach is well suited to identify these independent (often multidiisease) association signals (see “Stepwise subset-based conditional regression”).

The subset-based meta-analysis is a generalized fixed-effects meta-analysis and explores all possible subsets of diseases (for a restricted disease set if specified) for the presence of true association signals, while adjusting for the multipletesting required and a fixed control group shared by all diseases. To control for potential population stratification, we adjusted association test statistics by means of PCA using the top seven axes of variation (Supplementary Table 16). Adjusted two-tailed P_{SBM} values (risk versus protective) were obtained using the discrete local maxima (DLM) method estimating tail probabilities of the z-score test statistic that is maximized over a grid of neighboring subsets. The maximum (in absolute value) of the subset-specific z statistics is a conservative variable selector in the sense that, for large samples, it will select only non-null studies, but it is not guaranteed to select all of the non-null studies.

The genomic inflation factor (λ) calculated using 1,820 ‘null’ SNPs (outside the MHC region) associated with reading and writing ability, psychosis and schizophrenia was 1.082 (λ_{1,000} for an equivalent study of 1,000 cases and 1,000 controls = 1.002), indicating minimal evidence of residual population stratification in the overall data set of 52,262 cases and 34,213 controls. Where a particular SNP is only associated with a single disease, the standard meta-analysis methodology has slightly higher power than the subset-based approach. To avoid loss of statistical power in such settings, we looked up every SNP with P_{SBM} < 5 × 10^{-7} within and outside the 166 non-MHC susceptibility loci, to see whether genome-wide significance (P_{diagnosis} < 5 × 10^{-8}) was achieved in any of the five single disease versus control subgroups. Univariate association statistics (restricted to single-disease data sets versus the fixed control group) were obtained using the same DLM method. The increased statistical power of the single association test (P_{diagnosis}) in comparison to original individual-disease Immunochip analyses is likely due to the fact that the Immunochip data here with a larger sample size (except for ankylosing spondylitis) would be used as a screening tool instead of as a replication data set after screening smaller-sized GWAS discovery data sets. The large number of new ankylosing spondylitis loci could largely be attributed to an approximately 2.5x increase in the size of the control cohort as compared to the original ankylosing spondylitis Immunochip study (15,578 controls in the original study versus 34,213 controls in the current study). After ‘subtracting’ the new trans-ancestry Crohn’s disease and ulcerative colitis loci from the IBD trans-ancestry study, 27 of 35 new genome-wide significant non-MHC risk loci remain for the five diseases under study. Using an alternative method, we identified more pleiotropic loci shared by ulcerative colitis and Crohn’s disease (Supplementary Table 17 and Supplementary Note).

Stepwise subset-based conditional logistic regression. Lead (independent) SNPs associated with single and multiple diseases were selected through...
stepwise regression to condition away lead SNPs one at a time until no associations remained, following a recently published stepwise conditional SBM fine-mapping approach. This is an effective method for separating independent signals and assumes that LD between the independent causal variants is low. Significance was defined by Bonferroni correction of the number of LD-independent markers on the Immunochip (0.05/37,377 = 1.34 × 10^{-6}).

Cluster plot inspection. Immunochip intensity cluster plots of all genome-wide significant SNP markers (P_{SBM} < 5 × 10^{-7} and P_{diabetes} < 5 × 10^{-8}; P_{PSC} < 5 × 10^{-8}) from Supplementary Tables 2 and 3 were manually inspected by three different persons using Evoker to ensure that they were well clustered.

Bayesian multinomial regression for model selection. To compare different disease models at each locus, we used Bayesian multinomial regression. A disease model is a list of the diseases with which a given locus is associated (has a nonzero log-transformed odds ratio; for example, ‘associated with Crohn’s disease and ulcerative colitis but not with psoriasis, ankylosing spondylitis or PSC’ is one disease model, as is ‘associated with all diseases’). There are a total of 32 possible disease models for the five phenotypes, which includes the null model (‘not associated with any disease’). Our aim is to infer the posterior probability for each of these disease models, conditional on the genotype and phenotype data we have. We do this under a Bayesian setting, as it naturally handles the different uncertainties on the effect sizes for each disease due to their different sample sizes and powers. The methods we describe below are implemented in the open source Trinculo software package.

The Bayesian multinomial logistic regression software calculates a marginal likelihood for each model, incorporating uncertainty in the effect size as

\[ Pr(D | M) = \int Pr(D | \beta) Pr(\beta | M) d\beta \]

where \( \beta \) is a vector of log-transformed odds ratios for each disease \( D \). The likelihood \( Pr(D | M) \) is given by the multinomial logistic likelihood. The prior distribution on the effect sizes is given by \( \beta M - \text{MVN}(0, \Sigma M) \), where \( \Sigma M \) is the prior covariance matrix for model \( M \). To enforce phenotypes that are not associated with the disease, we set \( \sum_{ij}^M \beta_{ij} = 0 \) if either phenotype \( i \) or \( j \) is not associated with the locus. We use Newton’s method to calculate the maximum a posteriori estimate (MAP) for the parameters and calculate the marginal likelihood using a Laplace approximation around the MAP.

We calculate the posterior probability for each model as

\[ Pr(M | D) = \frac{Pr(D | M) Pr(M)}{\sum_M Pr(D | M') Pr(M')} \]

where \( Pr(M) \) is a per-model prior.

The method thus requires two priors: the model covariance matrices \( \Sigma M \) and the per-model priors \( Pr(M) \). We analyze each variant using six different sets of priors, two different forms of the covariance matrix prior and three different forms of the per-model prior. For the covariance prior, we use (i) a simple uniform prior across all models, (ii) a single covariance matrix where \( Pr(\Sigma M | D) = \sigma^{\text{MAP}} \), and (iii) an empirical covariance prior where the number of phenotypes associated with the locus, inferred by maximum likelihood.

We calculated posteriors for each model for each risk variant under the six different priors. For each risk variant, we then took a vote of the highest posterior models under each prior, such that we select whichever model was considered best by the largest number of priors. We also recorded how many priors voted for that model and how much posterior each prior gave to the winning model.

If SNPs represented secondary independent association signals identified from stepwise conditional SBM analysis, then we tested conditional on all other identified genome-wide significant independent signals within the same locus. Within the Bayesian logistic regression, we included the genotype at the lead SNP (and further preceding independent signals) as a covariate in the model.

Disease correlation measure and temporal comorbidity. To determine significant temporal co-occurrences (disease pairs) for the five inflammatory diseases under study, we screened an independent data set covering ICD10 diagnosis codes from 6,631,920 people of the entire Danish population in the period from 1996 to 2014 (ref. 5). We used relative risk to measure the strength of the correlation between a pair of diagnoses (diagnosis A followed by diagnosis B). Relative risk estimates and associated P values were calculated using a sampling approach as described in the original study. In brief, given a pair, diagnosis A followed by diagnosis B, relative risk of a temporal association was calculated as the ratio of the observed number of patients who had disease A then disease B within 5 years and the number of randomly matched control patients diagnosed with disease B within 5 years of a matched discharge. Each matched control was of the same age (birth decade) and sex as the case and had a discharge of the same type (inpatient, outpatient or emergency room) within the same calendar week as the case diagnosed with disease A (from which the 5 years to develop disease B was calculated). The significance threshold of \( P = 0.05/823,606 = 1.21 \times 10^{-8} \) was applied using Bonferroni correction for the testing of 823,606 directed pairs in the original study.

Distinguishing pleiotropy and heterogeneity. We used BUHMBOX v0.38 (Breaking Up Heterogeneous Mixture Based On Cross-Locus Correlations) to evaluate whether the sharing of risk alleles observed across pairs of diseases (disease A and disease B) was driven by true pleiotropy (where there is pervasive sharing of risk alleles by two diseases) or by heterogeneity (where a subgroup of disease A cases has a higher burden of risk alleles for disease B). The BUHMBOX approach has been described in detail elsewhere. Briefly, a GRS approach is used to detect significant sharing of risk loci by disease A and disease B. If such genetic sharing is detected using GRSs, the BUHMBOX test statistic (which identifies heterogeneity by calculating the cross-locus correlation of disease B–associated loci among disease A cases) is applied to verify whether these associations are due to heterogeneity (for example, sample misdiagnosis or excessive comorbidity) as opposed to biological pleiotropy. In the setting of pleiotropy, pleotropic disease B risk alleles are shared across all disease A cases, whereas in heterogeneity only a subset of disease A cases share disease B risk alleles. Thus, cross-locus correlations between disease B–associated loci are positive in the presence of heterogeneity but not in the case of pleiotropy. To strictly control for false positives, BUHMBOX uses LD pruning, the top seven principal components from PCA and delta-correlations between cases and controls.

First, to quantify pleiotropy for each of the 20 possible pairs of five diseases, we calculated GRSs using known independent risk loci for disease B for each case and control in the disease A sample (based on disease B risk alleles, weighted by effect size) and tested the association of these GRSs with disease A. We obtained the list of known associated loci from the previous literature for ankylosing spondylitis, Crohn’s disease, psoriasis, PSC and ulcerative colitis.

Next, we evaluated the presence of heterogeneity by applying BUHMBOX to each of the 20 pairs of diseases. We estimated the statistical power of BUHMBOX to detect a certain proportion of sample heterogeneity by simulation (Supplementary Fig. 10 and Supplementary Table 13), using the effect sizes and allele frequencies of the disease B loci and randomly simulating the number of cases and controls in the disease A sample.

Functional annotation of associated variants. The variants identified in this study were annotated using the Ensembl variant effect predictor (VEP).
(release 77) to determine genomic position annotations, including the closest gene, and functional consequences (using the most severe consequence from SIFT\textsuperscript{52} and PolyPhen\textsuperscript{53}). The --assembly flag was set to GRCh37 and the --pick flag was added to retrieve the most severe consequence for the variants. The UpDownDistance plugin was used to retrieve the nearest gene ID within 10 kb of the variant. TSS distance was retrieved using the TSSDistance plugin. We also included the --regulatory flag to annotate where a variant overlaps a regulatory feature. The DNase hypersensitivity site (DHS) and promoter annotations were taken from 1000 Genomes Project annotations\textsuperscript{54}.

To determine whether any of the lead variants were in high LD ($r^2 > 0.8$) with a functional variant, the 1000 Genomes Project v3 EUR haplotypes were used (1000 Genomes Phase III 20130502 release). Pairwise LD was calculated between the lead SNPs and all other SNPs in this data set using PLINK (v1.09)\textsuperscript{55}. Only variants that occurred in the 1000 Genomes Project data set were included in this analysis. The GWAS catalog\textsuperscript{19} was used to determine whether any lead variants or variants in high LD ($r^2 > 0.8$) with the lead variants had previously been reported in other GWAS. ImmunoBase and Europe PubMed Central were also used to determine whether variants had previously been associated with an autoimmune phenotype.

40. Trynka, G. \textit{et al.} Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. \textit{Nat. Genet.} \textbf{43}, 1193–1201 (2011).

41. 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. \textit{Nature} \textbf{491}, 56–65 (2012).

42. Purcell, S. \textit{et al.} PLINK: a tool set for whole-genome association and population-based linkage analyses. \textit{Am. J. Hum. Genet.} \textbf{81}, 559–575 (2007).

43. Price, A.L. \textit{et al.} Principal components analysis corrects for stratification in genome-wide association studies. \textit{Nat. Genet.} \textbf{38}, 904–909 (2006).

44. International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. \textit{Nature} \textbf{449}, 851–861 (2007).

45. Price, A.L. \textit{et al.} Long-range LD can confound genome scans in admixed populations. \textit{Am. J. Hum. Genet.} \textbf{83}, 132–135, author reply 135–139 (2008).

46. Abraham, G. & Inouye, M. Fast principal component analysis of large-scale genome-wide data. \textit{PLoS One} \textbf{9}, e93766 (2014).

47. Taylor, J.E., Worsley, K.J. & Gosselin, F. Maxima of discretely sampled random fields, with an application to ‘bubbles’. \textit{Biometrika} \textbf{94}, 1–18 (2007).

48. Tsoi, L.C. \textit{et al.} Enhanced meta-analysis and replication studies identify five new psoriasis susceptibility loci. \textit{Nat. Commun.} \textbf{6}, 7001 (2015).

49. Liu, J.Z. \textit{et al.} Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. \textit{Nat. Genet.} \textbf{47}, 979–986 (2015).

50. Morris, J.A., Randall, J.C., Maller, J.B. & Barrett, J.C. Evoker: a visualization tool for genotype intensity data. \textit{Bioinformatics} \textbf{26}, 1786–1787 (2010).

51. McLaren, W. \textit{et al.} Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. \textit{Bioinformatics} \textbf{26}, 2069–2070 (2010).

52. Ng, P.C. & Henikoff, S. SIFT: predicting amino acid changes that affect protein function. \textit{Nucleic Acids Res.} \textbf{31}, 3812–3814 (2003).

53. Adzhubei, I.A. \textit{et al.} A method and server for predicting damaging missense mutations. \textit{Nat. Methods} \textbf{7}, 248–249 (2010).

54. Gusev, A. \textit{et al.} Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. \textit{Am. J. Hum. Genet.} \textbf{95}, 535–552 (2014).

55. Chang, C.C. \textit{et al.} Second-generation PLINK: rising to the challenge of larger and richer datasets. \textit{Gigascience} \textbf{4}, 7 (2015).