

## **Rare Comorbidity between Inflammatory Bowel Disease and Primary Biliary Cholangitis:**

### **Evidence from Causality, Shared Genetic Architecture and Transcriptomics**

Wentao Huang<sup>a,b†</sup>, Rui Jiang<sup>a,c†</sup>, Ruijie Zeng<sup>a,d†</sup>, Yuying Ma<sup>a,b</sup>, Lijun Zhang<sup>a,c</sup>, Shuangshuang Tong<sup>a,d</sup>, Yanlin Lyu<sup>a,d</sup>, Jiaxuan Wang<sup>a</sup>, Felix W Leung<sup>f,g\*</sup>, Weihong Sha<sup>a,b,c,d\*</sup>, Hao Chen<sup>a,b,c,d\*</sup>

<sup>a</sup>Department of Gastroenterology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou 510080, China

<sup>b</sup>The Second School of Clinical Medicine, Southern Medical University, Guangzhou 510006, China

<sup>c</sup>School of Medicine, South China University of Technology, Guangzhou 510006, China

<sup>d</sup>Shantou University Medical College, Shantou 515041, China

<sup>e</sup>Faculty of Synthetic Biology, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

<sup>f</sup>Sepulveda Ambulatory Care Center, VA Greater Los Angeles Healthcare System, Los Angeles 91343, California, USA.

<sup>g</sup>University of California Los Angeles David Geffen School of Medicine, Los Angeles 90095, California, USA.

#### **\*Corresponding authors:**

Prof. Felix W Leung, University of California Los Angeles David Geffen School of Medicine, Los Angeles 90095, California, USA.

E-mail: [felixleung@socal.rr.com](mailto:felixleung@socal.rr.com)

Prof. Weihong Sha, Department of Gastroenterology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou 510080, China;

E-mail: [shaweihong@gdph.org.cn](mailto:shaweihong@gdph.org.cn)

Prof. Hao Chen, Department of Gastroenterology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou 510080, China;

E-mail: [chenhao@gdph.org.cn](mailto:chenhao@gdph.org.cn)

**†These authors contributed equally to this work**

## Abstract

**Background:** Clinical studies have found comorbidity between Inflammatory Bowel Disease (IBD) and primary sclerosing cholangitis (PSC). Primary biliary cholangitis (PBC) is another autoimmune liver disease but the coexistence of IBD and PBC is rare. Whether there exists comorbidity between IBD and PBC and potential mechanism remains unclear.

**Methods:** We assessed the casual effect between PBC and IBD, i.e., Crohn Disease (CD) and Ulcerative Colitis (UC) independently based on genome-wide association studies (GWAS) summary statistics. By leveraging data from GWAS data, Bulk tissue RNA sequencing (bulk RNA-seq) data, and Single-cell RNA sequencing (scRNA-seq) dataset, we investigated the shared genetic architecture between IBDs and PBC. The transcriptomic expressions of shared genes were explored in patients with IBD (intestinal biopsies) and PBC (peripheral CD4<sup>+</sup> T cells).

**Result:** We found a bidirectional causal relationship for PBC and IBDs using Mendelian randomization. The IBDs had been considered as the protective factors on PBC (0.87[95% confidence interval (CI): 0.81-0.93],  $P = 8.72e-5$ , vice versa (0.91[95% CI: 0.81-0.93],  $P = 2.65e-09$ ). We find a consistent negative genetic correlation between PBC and IBD (LDSC:  $r_g = -0.2245$ ,  $P = 2.89e-5$ ). Cross-trait analysis yielded 9 shared risk SNPs and 7 nearest genes. In transcriptome analysis, we observed significant ( $P < 0.05$ ) differences expression in intestinal biopsies (*PGAP3* and *DENND1B*) and in peripheral CD4<sup>+</sup> T cells (*PTPN11* and *PNMT*). We identified shared tissue-specific heritability enrichment for PBC and IBD (including CD not UC) in lung, spleen and cells EBV-transformed lymphocytes and identified shared cell type-level enrichment for IBD, CD and PBC in type 1 dendritic cells, natural killer cells, CD8<sup>+</sup> cytotoxic T lymphocytes in lung and activated CD8<sup>+</sup> T cell in spleen.

**Conclusion:** Our study indicates that IBD and PBC are protective factors for each other and shared genetic architecture may contribute to the negative genetic correlation. These findings may explain the rare comorbidity between IBD and PBC.

**Keywords:** Inflammatory Bowel Diseases, Primary Biliary Cholangitis, Causality, Shared Genetic Architecture, Comorbidity

## Introduction

Inflammatory bowel disease (IBD) is a group of the chronic immune-mediated idiopathic inflammation of the gastrointestinal tract and has two forms including Crohn's Disease (CD) and Ulcerative Colitis (UC)<sup>1-4</sup>. IBD always experience various hepatobiliary disorders, including autoimmune liver diseases<sup>5</sup>. As one of autoimmune liver diseases, 70%-80% patients with primary sclerosing cholangitis (PSC) would have IBD and share a similar or common pathogenesis have been hypothesized<sup>6</sup>. Primary biliary cholangitis (PBC) is another autoimmune liver disease and causes chronic and persistent cholestasis in the liver, eventually resulting in cholangitis and hepatic failure without appropriate treatment<sup>7,8</sup>. PBC was always accompanying with various extrahepatic autoimmune diseases<sup>9</sup>. However, unlike PSC, PBC has little associated with IBD<sup>5</sup> and there are few reports on the coexistence of PBC and IBD<sup>10</sup>. Patients with PBC were diagnosed after IBD had been diagnosed<sup>11</sup> and coexistence with IBD had no impact on PBC outcome<sup>12</sup>. In addition, during rare comorbidities reports, PBC have higher rates of coexistence with CD than UC<sup>13-15</sup>. Whether there existed the causality and absence of comorbidities between PBC and CD or UC remain unknown.

Prior studies have revealed that genetic risk factors lead to PBC<sup>16</sup> and IBD<sup>17,18</sup>. Genome-wide association study (GWAS) analysis had revealed strong genetic effect in the major histocompatibility complex (MHC)<sup>19,20</sup> and non-MHC<sup>19,21-23</sup> both PBC and IBD. These loci may contribute to the abnormal immune activation<sup>24</sup> for PBC and was associated with IBD<sup>19</sup>. Further investigation identified tissue-specific and cell-type-specific enrichment for IBD in CD4<sup>+</sup> T cells in lung and CD8<sup>+</sup> cytotoxic T cells<sup>25,26</sup>, which had a significant role in IBD and PBC pathogenesis<sup>27</sup> including inflammation<sup>28-30</sup> and immune activation<sup>31,32</sup>. These findings suggest that genetic factors can influence the rare coexistence of PBC and IBD but the SNP heritability enrichment for PBC remains unclear. Thus, further evaluating the shared genetic architecture may contribute to understand the potential mechanism of the rare comorbidity and explore new therapeutic targets for a precision medicine.

In the current study, based on Single-cell RNA sequencing (scRNA-seq) dataset, GWAS dataset and Genotype-Tissue Expression (GTEx) dataset, five studies were performed to dissect: (1) causal relationship between IBD and PBC relied on five Mendelian randomization methods; (2) genetic correlation between IBD and PBC; (3) shared SNP combining Heritability Estimation

from Summary Statistics(p-HESS) and cross-trait GWAS meta-analysis; (4) transcriptomic evaluation of the risk genes; and (5) SNP heritability enrichment in the specific tissue and cell-type.

## **Methods**

### **Data samples**

#### **GWAS summary statistics**

GWAS summary results for IBD conclude three traits (IBD,CD and UC)<sup>33</sup>. GWAS summary statistics for IBD include 34,915 controls and 25,042 cases of European ancestry. The data contain 40,266 Europeans participant including 28,072 controls and 12,194 cases for CD and 45,975 Europeans Europeans participant including 33,609 controls and 12,366 cases for UC. GWAS summary statistics for PBC download from the public GWAS summary statistics, which contained 8,021 European ancestry cases and 16,489 European ancestry controls<sup>34</sup>.The datasets used for replication are another public PBC GWAS summary statics from the EBI database (including 2,764 cases and 10,475 controls<sup>35</sup>). GWAS summary statistics were handled with the collection of samples and the quality control procedures, which was described in the publication<sup>26,36</sup>.

#### **Bulk-tissue RNA sequencing gene expression data**

The bulk-tissue RNA-seq gene expression data was available at GTEx project, contains gene expression of 53 non-diseased human primary tissues<sup>37</sup>, to perform the subsequent linkage disequilibrium score regression specifically expressed genes (LDSC-SEG). We chose the GTEx v6p dataset, which has been fixed<sup>25</sup>. For transcriptome analysis, bulk RNA-seq for IBD, CD, UC and PBC were from the Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) to evaluate the gene expression of shared risk genes in each trait. The GSE179285 dataset<sup>38</sup>, which collected intestinal biopsies from 254 samples that corresponded to 168 CD patients, 55 UC patients and 31 normal subjects. The gene expression matrix of IBD was a combination of CD and UC patients (N = 223). The GSE93170 dataset contained a transcriptome-wide analysis of peripheral CD4<sup>+</sup> T cells derived from 6 PBC cases and 6 control subjects<sup>39</sup>.

### **Single-cell RNA sequencing gene expression data**

Based on evidence of enrichment in heritability of tissue-level SNPs in LDSC-SEG analysis, we matched the publicly available single-cell scRNA-seq data for four enrichment tissues including spleen (N= 94,257 immune cells)<sup>40</sup>, lung (N = 57,020 cells)<sup>40</sup>, small-intestine terminal ileum (N= 53,193 epithelial cells from mouse)<sup>41</sup> and hematopoietic stem and progenitor cell (HSPC, N= 7,551 adult human blood cells)<sup>42</sup>. For small-intestine terminal ileum data, the gene symbols were converted from mouse to human using the “EWEC” package. In total, there are 81 cell types from four tissues were included.

### **Statistical analyses**

#### **Mendelian Randomization Analyses**

We selected five MR methods including Generalized Summary-data-based Mendelian Randomization (GSMR)<sup>43</sup>, weighted median<sup>44</sup>, MR-Egger<sup>45</sup>, weighted mode<sup>46</sup>, and inverse variance weighting (IVW)<sup>47</sup> to test the causality between PBC and IBD (including UC and CD), from the R packages “TwoSample MR”<sup>48</sup> and “GSMR”<sup>43</sup> ( $P < 0.05$ ) To validate the stability of the causal relationship, we also conducted a validation study using the validation group.

We performed sensitivity analysis including Leave-one-out sensitivity test, pleiotropy test, and heterogeneity test. MR-Egger regression<sup>49</sup> was used to assess whether the MR-Egger intercept was different from zero and the pleiotropy of instruments<sup>45</sup>. The heterogeneity was assessed by Cochran’s Q statistic<sup>50,51</sup>. Outlier instrument was removed by leave-one-out analysis<sup>52,53</sup> and MR-PRESSO analysis<sup>49</sup>. In MR-PRESSO analysis, it tests those SNPs cause the heterogeneity in the estimate of the causality and estimated the global test  $P$  value after excluding these SNPs. Leave-one-out analysis was utilized to test whether there existed special single variant influencing the causality, which was remained and had combined effect value. We also calculated the causal effect on two traits using GSMR method. HEIDI-outlier test would remove those SNPs that show evidence of pleiotropic effects by the heterogeneity (HEIDI-outlier test  $< 0.05$ ) were removed<sup>43</sup>. After these methods, those SNPs which show genome-wide significance ( $P < 5 \times 10^{-8}$ ) from the ‘exposure’ trait was set as the instrumental variables. The validation group also performed the sensitivity analysis.

## **Heritability and Genetic Correlation**

LD score regression (LDSC)<sup>54</sup> (Python 2.7) was performed to evaluate the genetic correlation for different subtypes. Those SNPs which did not match the reference panel (INFO score  $\leq 0.9$  or  $MAF \leq 0.01$ ) would be excluded based on the LD scores that are computed from the 1000 Genomes project phase III<sup>55</sup>. Based on stratified linkage disequilibrium score regression (SLDSC) with the baseline-LD model<sup>2</sup>, we calculated single trait SNP heritability for IBDs and PBC. Then bivariate LDSC<sup>56</sup> were conducted to evaluate genetic correlations between IBDs and PBC<sup>57</sup>.

We also performed Genetic covariance analyzer (GNOVA)<sup>58</sup> to calculate the genetic correlation and the SNP-based heritability for IBDs and PBC. GNOVA explore genetic covariance using all genetic variants shared between two diseases. We then calculated genetic correlation based on genetic covariance and variant heritability. In addition, sample overlap correction of GWAS summary statistics were corrected.

## **Evaluating of local genetic correlations using $\rho$ -HESS**

$\rho$ -HESS<sup>59</sup> is utilized to examine local SNP-heritability and genetic correlation from GWAS summary data. To identify whether there is shared genetic correlation between IBDs and PBC at the local independent region in the genome, we calculated the local genetic correlations using  $\rho$ -HESS (Python 2.7). 1,699 regions were approximately LD-independent loci with average size of nearly 1.5Mb<sup>60</sup>. Finally, relied on the 1000 Genomes project as reference, we estimated genetic correlation between two subtypes and the local SNP heritability for each traits<sup>61</sup>.

## **Cross-trait meta-analysis**

We conducted two independent Cross-trait meta-analysis of GWAS summary statistics including cross-phenotype association test (CPASSOC)<sup>62</sup> and multi-trait analysis of GWAS (MTAG)<sup>63</sup> to identify shared risk SNPs between PBC and each of IBD, CD and UC. Based on a strong genetic correlation between two traits, MTAG performs meta-analysis of different traits when all variants have the same genetic correlation across all traits. The results of MTAG are effective with possible sample overlap between GWAS summary statistics for traits. We assessed more than 10% of SNPs were causal for each trait by calculating the value of maxFDR. CPASSOC analysis is suitable for

binary traits but needs to account for potential false positives due to sample overlap. Shared risk SNPs were determined to be significant ( $P < 5 \times 10^{-8}$ ) in both CPASSOC and MTAG. These independent SNPs were identified by LD clumping ( $r^2 \leq 0.001$  within 10,000-kb windows) in Plink. Eventually, shared risk SNPs were identified significant if they also surpassed Bonferroni-correction ( $N = 1,562$ ,  $P < 3.2 \times 10^{-5}$ ) for local genetic correlation in  $\rho$ -HESS analysis.

### **Transcriptomic evaluation of the risk genes**

We evaluated and compared whether the shared risk genes obtained from the above analysis exhibited significant ( $P < 0.05$ ) differences in genes expression occurred in IBD, CD, UC, PBC patients and normal subjects. We selected the GSE179285 dataset<sup>38</sup>, which is a collection of 254 samples that corresponded to 168 CD patients, 55 UC patients and 31 normal subjects. The gene expression matrix of IBD was a combination of CD and UC patients ( $N = 223$ ). The GSE93170 dataset contained a transcriptome-wide analysis derived from 6 PBC cases and 6 control subjects. The “ggplot2” and “ggpubr” packages were used for the estimation of differences among the gene expression profiles and visualization.

### **LD score regression applied to specifically expressed genes (LDSC-SEG)**

We implemented LDSC-SEG<sup>25</sup> to infer the enrichment for SNP heritability and the related tissue-specific enrichment in IBD, CD, UC and PBC, using specifically expressed genes sets in tissues and GAWSS summary statistics for each trait. We utilized 1000 Genomes Project in Phase 3 as the reference panel and only HapMap 3 SNPs with minor allele frequency (MAF)  $> 0.05$  were included. The gene sets utilized in LDSC-SEG were defined by Finucane et al. for the genotype tissue expression (GTEx) project with 53 tissue types. According to the t-statistic, the top 10% of candidate genes were selected as the set of genes specifically expressed in tissues. The regression coefficient p-values were determined from the coefficient Z-score, and the false discovery rate (FDR)  $< 0.05$  was retained to represent the significance of enrichment tissues across PBC and each trait of IBD, CD and UC.

### **Multi-marker Analysis of Genomic Annotation (MAGMA)**

Cell type-specific enrichment was performed by MAGMA method<sup>64</sup>, which converted SNP

associations into gene-level associations and estimated enriched cell types via differential expression data from the input scRNA-Seq data. We matched the publicly available single-cell scRNA-seq data for four tissues including spleen (N= 94,257 immune cells)<sup>40</sup>, lung (N = 57,020 cells)<sup>40</sup>, small-intestine terminal ileum (N= 53,193 epithelial cells from mouse)<sup>41</sup> and HSPC (N= 7,551 adult human blood cells, GSE137864)<sup>42</sup> which were enriched in LDSC-SEG. MAGMA enrichment results were analyzed with “MAGMA.Celltyping” R package (version 2.0.7) and all scRNA-Seq data for tissues were processed with the “EWEC” R package (version 1.5.8)<sup>65</sup>. Results with  $FDR > 5 \times 10^{-3}$  indicated significant and independent genetic signals relative to cell type.

## Results

### Causality between IBD and PBC

The bi-directional MR were conducted to assessed the casual association between IBD and PBC. Our results found the causal effect of IBD on PBC in five methods (0.87[95% CI: 0.81-0.93],  $P = 8.72e-5$ ) (**Fig. 1 and Supplementary S4**). IBD played the protective factor to the PBC. No significant horizontal pleiotropy for the causality of IBD on PBC existed (MR-PRESSO global test  $P = 0.089$ ; MR-Egger intercept =  $3.56e-4$ , SE = 0.013,  $P = 0.98$ ; **Supplementary Fig.S1**).

In the reverse analyses, five methods found consistent evidence for causality of PBC on IBD (**Fig.1 and Supplementary S4**). PBC was also been considered as the protective factor to the IBD (0.91[95% CI: 0.81-0.93],  $P = 2.65e-09$ ). IBD played the protective factor to the PBC. There was no horizontal pleiotropy for the causality of IBD on PBC (MR-PRESSO global test  $P = 0.09$ ; MR-Egger intercept = -0.0178, SE = 0.0101,  $P = 0.094$ ; **Supplementary Fig.S5**). The causal association was verified using the validation sets (**Supplementary Fig.S8**).

### Causality between PBC and two subtypes of IBD

We further explored casual effect for two phenotypes of IBD (CD and UC) on PBC. There existed the causality of CD and UC on PBC. (CD: 0.92 [95% CI: 0.88-0.96],  $P = 4.23e-4$ ; UC: 0.89[95% CI: 0.81-0.97],  $P = 0.00985484$ ) (**Fig. 1 and Supplementary Fig.S4**). Both two subtypes played the protective factor to the PBC. No significant horizontal pleiotropy for the causality of CD on PBC was tested (CD: MR-PRESSO global test  $P = 0.218$ ; MR-Egger intercept = -0.0025, SE =



0.011,  $P = 0.82$ ; UC: MR-PRESSO global test  $P = 0.223$ ; MR-Egger intercept  $= -0.052$ , SE = 0.02742005,  $P = 0.07$ ; **Supplementary Fig.S2-S3**).

In the reverse analyses, five methods found consistent evidence for causal effect of PBC on CD and UC (**Fig.1 and Supplementary S4**). PBC was also been considered as the protective factor to the two subtypes (CD: 0.88[95% CI:0.85-0.91],  $P = 4.76e-11$ ; UC: 0.88[95% CI: 0.85-0.91],  $P = 4.76e-11$ ). CD and UC played the protective factor to the PBC. There was not significant horizontal pleiotropy for the causality of CD and UC on PBC (CD: MR-PRESSO global test  $P = 0.313$ ; MR-Egger intercept =  $-0.022$ , SE = 0.011,  $P = 0.05$ ; UC: MR-PRESSO global test  $P = 0.09$ ; MR-Egger intercept  $= -0.0033$ , SE = 0.012,  $P = 0.79$ ; **Supplementary Fig.S6-S7**). The causal effect was verified using the validation sets (**Supplementary Fig.S8**).

#### **Estimation of Genetic correlations between IBDs and PBC**

We calculated the SNP heritability for IBDs and PBC using LDSC<sup>54</sup> with the baseline-LD model<sup>66</sup>. the SNP heritability for IBDs<sup>26</sup> and for PBC was consistent with the previous study. The results of bivariate LDSC show that IBD including CD and UC have negative correlation with PBC (**Supplementary Table 1**). To get a robust result, we conducted GNOVA and explored a consistent result between PBC and IBDs after correcting sample overlap (**Supplementary Table 1**).

#### **Local genetic correlations between IBD, UC, CD and PBC**

Local genetic correlation was evaluated for 1,561 genomic partitions and Bonferroni significant regions were assessed (**Supplementary Table S1-S3**). Analysis of local genetic correlation revealed loci of significant local genetic covariance to four traits (**Supplementary Fig. S9-S11**). Polygenicity was higher for PBC than for IBDs. There was close agreement in the average local genetic correlation in regions harboring PBC-specific loci or IBDs-specific loci (**Fig. 2**). We estimated the local single-trait SNP heritability for PBC and IBDs (PBC: 0.255, IBD:0.466, CD:0.497, UC:0.296). Genome-wide local genetic correlations tested by HESS were all largely consistent with genetic correlation estimated by LDSC and GNOVA (IBD-PBC: -1.12, CD-PBC: -1, UC-PBC: -1.3) between IBDs and PBC. Polygenicity was higher for IBDs than for PBC

(**Supplementary Fig.S12**). The potential correlation for sharing of genetic variation may exist in the whole genome.

### **Shared risk SNPs were identified between PBC and each trait of IBD, CD and UC**

With significant genetic correlations between PBC and each of IBD, CD and UC, we utilized two complementary cross-trait meta-analysis approaches to identify shared risk SNPs between IBDs and PBC. With significant local genetic correlation, we identified that six SNPs (rs11583319, rs7583409, rs6580223, rs11066301, rs407307, rs5757587) were significantly ( $P < 5 \times 10^{-8}$ ) associated with the joint phenotype IBD-PBC, three SNPs (rs67236816, rs407307, rs5757587) were significant in the CD-PBC cross-trait GWAS, three SNPs (rs10752745, rs12946510, rs5757587) were significant in the UC-PBC. We totally obtained 9 candidate shared risk SNPs including two previously discovered SNPs (rs6580223 and rs12946510) associated with IBD and/or PBC and seven potential novel SNPs. The maxFDR values for MTAG analyses of PBC and each of IBD, UC, and CD were 0.014, 0.005, 0.004 and 0.026 respectively. The lists of shared risk SNPs were summarized in **Table 2**.

### **Differential expression of the risk genes and correlation of gene expressions**

We further evaluated the mRNA expressions of candidate risk genes (*DENND1B*, *DNMT3A*, *NDFIP1*, *PTPN11*, *PNMT*, *PGAP3*) in PBC and each of IBD, CD and UC patients. Significant ( $P < 0.05$ ) differences of *PGAP3* expression and *DENND1B* expression were observed in patients with UC ( $P_{PGAP3} = 0.013$ , **Supplementary Fig. S13C**;  $P_{DENND1B} = 1.5 \times 10^{-3}$ , **Supplementary Fig. S13D**) and control subjects. The expression of *PGAP3* was significantly lower than that of the control in IBD samples ( $P = 0.021$ , **Supplementary Fig. S13A**), CD patients ( $P = 0.035$ , **Supplementary Fig. S13B**) and UC samples ( $P = 0.013$ ; **Supplementary Fig. S13C**). In PBC patients, the expression of *PTPN11* was significantly higher compared with the control group ( $P = 8.7 \times 10^{-3}$ , **Supplementary Fig. S13E**), and conversely, the expression of *PNMT* was decreased ( $P = 0.041$ , **Supplementary Fig. S13F**).

### **SNP heritability enrichment of SNPs in IBD, CD, UC and PBC-related tissues**

We used LDSC-SEG to investigate tissue-specific enrichment of SNP heritability for IBD, UC, CD and PBC. We detected a significant enrichment at an FDR > 0.05 threshold for tissues, IBD and CD exhibited FDR-significant enrichment in lung, spleen, whole blood and small

intestine-terminal ileum tissues, CD also enriched in cells EBV-transformed lymphocytes. For PBC, we observed significant enrichment in three tissues including lung, spleen and cells EBV-transformed lymphocytes. Lung and spleen were commonly associated tissues with IBD, CD and PBC (**Figure 3A-D**). However, there was no significant evidence for a specific enriched tissue for UC. The magnitude of SNP heritability enrichment in these tissues are presented in **Supplementary Table S4 and Supplementary Figure S14**.

### **Cell type-level enrichment in IBD, CD, UC and PBC**

To identify the enrichment of cell types in spleen, lung tissues, we performed MAGMA to test for a correlation between cell type expression specificity and GWAS summary statistics of IBD, CD, UC and PBC. Of the enrich tissues identified in previous LDSC-SEG for IBD, CD, UC and PBC. We detected an enrichment at an  $FDR > 5 \times 10^{-3}$  threshold for enrichment tissues. In lung tissues, we observed enrichment in  $CD4^+$  T cells, type 2 dendritic cells (DC2), Monocyte cells, Nature killer (NK) cells,  $CD8^+$  cytotoxic T lymphocytes for IBD; and DC2,  $CD4^+$  T cells, NK cells, Monocyte cells, dendritic cells (DC1) DC1, regulatory T cells,  $CD8^+$  cytotoxic T lymphocytes for CD; and activated DC cells, DC1,  $CD8^+$  cytotoxic T lymphocytes, NK cells for PBC (**Supplementary Fig. S15**). In spleen tissues, NK\_FCGR3Apos, activated  $CD8^+$  T cells,  $CD8$  cytotoxic T lymphocytes were enriched in IBD; NK\_FCGR3Apos and activated  $CD8^+$  T cells were enriched in CD; CD 8 mucosa-associated invariant T cell, activated  $CD8^+$  T cells were enriched in PBC (**Supplementary Fig. S16**). We observed shared cell type-level enrichment for IBD, CD and PBC in DC1, NK cells,  $CD8$  cytotoxic T lymphocytes in lung and activated  $CD8^+$  T cells in spleen (**Supplementary Fig. 4A-4C**). None of the cell types in small-intestine terminal ileum and HSPC showed significant specificity in all four phenotypes. The results of cell-specific types show enrichment were described in detail in **Supplementary Table S5 and Supplementary Figure S15-S18**.

### **Discussion**

To our knowledge, we firstly confirmed mutual protective association and negative genetic correlation between IBD and PBC. Combining novel GWAS datasets, tissue and cell-type specific expression data, we explored the shared genetic architecture between IBD and PBC to explain the

rare comorbidity for two diseases.

In the previous study, the coexistence of both conditions is rare. A retrospective analysis (2006-2016), the largest series reported to date of PBC-associated IBD, reported six patients with PBC-associated IBD<sup>11</sup>. The number of cases reported by other clinicians is similar<sup>67-70</sup>. During the current study, based on the latest GWAS summary data, we firstly performed Mendelian randomization studies and identified a consistent bidirectional causality between IBD and PBC. The mutual protective roles may explain the rare comorbidity between IBD and PBC. These results provide the evidence that the genetic factors may contribute to the occurrence and development of comorbidity.

We identified the genetic correlation to provide evidence for the clinical arrange, drug discovery and disease prediction of comorbidity. The previous study found that the chromosome loss played an important role in the occurrence and the complication of PBC<sup>71,72</sup>. The previous GWAS study had identified chromosome 11q23.1<sup>73</sup> and 3q13.33<sup>74</sup> as susceptibility locus. These results may indicate that pleiotropic effects on the these traits exist in numerous genetic variants across the genome for a consistent  $r_g$  estimated by  $\rho$ -HESS, LDSC and GNOVA<sup>26</sup>.

To explore the shared risk SNPs between PBC and each of IBD, CD and UC, we combined HESS with joint analysis of GWAS summary results from related traits. We found two previously reported SNPs and seven novel SNPs with biologically plausible genetic associations with IBD-PBC, CD-PBC or UC-PBC. CD4<sup>+</sup> T cells play a significant role in immunological dysfunction observed in PBC<sup>75</sup> and IBDs<sup>76</sup>. The expression of nearest genes to the loci in CD4<sup>+</sup> T cells were associated with the inflammatory responses which may mediate the protective mechanism between PBC and IBDs. Polymorphisms in *NDFIPI* are associated with the development of IBD<sup>77</sup> and the critical function for *NDFIPI* in peripheral tolerance is as an induced, cell-autonomous brake against effector CD4<sup>+</sup> cells differentiation<sup>78</sup>. The expression of *PGAP3* reduced in noninflamed ileum biopsies from individuals in the overall IBD cohort<sup>79</sup>. The expression of *PGAP3* in CD4<sup>+</sup> cells of PBC is reduced at the transcriptome level. GWAS analysis found that noncoding polymorphisms in the de novo methyltransferase enzyme *DNMT3A* locus are associated with a higher risk for IBD<sup>80</sup>. The suppression of *DNMT3A* mRNA levels<sup>81</sup> have an effect on the genetic risk for IBD. Our results proved this point in transcriptome validation. The

previous study had indicated that the expression levels of *PTPN22*, *PTPN11*, and *PTPN2* play a significant role in actively inflamed intestinal tissue. Loss of *PTPN11* is also associated with increased colitis severity in vivo<sup>82</sup>. The downregulation of *PTPN11* would reduce the production of reactive oxygen species and protect intestine against oxidative stress-induced injury<sup>83</sup>. In PBC patients, the expression of *PTPN11* is also reduced at the transcriptome level. Therefore, the expression of the nearest gene to the risk SNP in CD4<sup>+</sup> T cells may contribute to the protective mechanism between PBC and IBDs.

In the other hand, the functional variant drives the nearest gene expression. Hitomi et al identified the rs12946510 affects the pathogenesis of PBC by altering *FOXO1* binding affinity in vitro to affect functional variation in gene expression<sup>84</sup>. Meanwhile, *FOXO1* plays a role in mucosal barrier injury of inflammatory bowel disease through TLR4/MyD88/MD2-NF-κB inflammatory signaling<sup>85</sup>. And down-regulation of *FOXO1* leads to over-production of Th17 cells and increased inflammatory responses<sup>86</sup>. *DENND1B* has been previously reported as one of the susceptibility gene regions that are associated with IBD(rs12740041)<sup>87</sup> especially CD and PBC( ( rs2488393 ) )<sup>88</sup>. However, our results found that *DENND1B* expression was significantly reduced in UC at the transcriptomic level. Because we identified a novel risk SNP ( rs11583319 ) to *DENND1B* and this may explain the difference. Four SNPs (rs67236816, rs407307, rs5757587, rs10752745) and three genes (*PNMT*, *PGAP3* and *LOC100996583*) that have never been reported to be genetically associated with PBC or IBD, including CD and UC. In our analysis, *PNMT* are notable for its depressed expression in PBC and elevated expression in IBD including UC and CD. These results indicated that the biological mechanisms of these novel SNPs and genes underlie disease susceptibility to PBC and each trait of IBD, CD and UC are required to be investigated by further functional research.

Although IBD and PBC primarily affect the small intestine and liver, respectively, there is growing evidence that other tissues and organs might also contribute to their pathogenesis. Tissue enrichment analysis illustrated that IBD, CD and PBC were mainly enriched in immune tissues such as lung, spleen and EBV-transformed lymphocytes. Lung and spleen presented co-enrichment in IBD, CD and PBC, but not UC, suggesting dissimilar shared etiologies between PBC-UC and PBC-CD in different tissues. IBD, including CD and UC, is characterized by chronic

inflammatory conditions in the intestinal tract, mainly resulted from a dysfunctional regulation of the immune system and an imbalance of the intestinal microbiota. Likewise, PBC is also characterized by the immune-mediated progressive cholestatic liver disease<sup>89</sup>. It has been reported that the possible immune pathways of IBD leading to lung lesions include dysfunctional immune cell homing, systemic inflammation, and microbial dysbiosis<sup>90</sup>. Changes in spleen size in patients with PBC may be useful to predict advanced fibrosis<sup>91</sup>. Previous studies have often examined tissue in the small intestine of IBD patients or liver of PBC patients. Our results suggest that IBD-PBC and CD-PBC might have a common pathogenic mechanism in common enriched tissues, such as lung and spleen. Although SNP enrichment for UC in tissue special levels was not significant, based on the strong genetic correlation and the risk SNP was identified, we suspected UC might have shared genetic overlaps with PBC, which deserves further study.

We then extended MAGMA to the single cell-level enrichment and identified three cell types (DC1, CD8<sup>+</sup> cytotoxic T lymphocytes, NK cells) in lung and activated CD8<sup>+</sup> T cell in spleen with significant heritability enrichments in IBD-PBC and CD-PBC. Inflammatory products induce DC cell maturation, and DC activation is a contributing factor to the development of IBD<sup>92,93</sup>. DC cells can induce a large number of effector T cells, and CD8<sup>+</sup> T cells are involved in the pathogenesis of IBD<sup>94</sup>. The decrease of CD8<sup>+</sup> T cell responses to commensal bacteria leading to Trm deficiency in the colon could lead to IBD<sup>95</sup>. Consistent with previous research results, DC cells, T cells and NK cells all played a role in the pathogenesis of PBC<sup>96</sup>. Activated T cells, particularly CD8<sup>+</sup> T cells, play a critical role in the destruction of biliary cells and mediate the pathogenesis of PBC in both human and murine models<sup>97</sup>. There existed no difference for the numbers of DC (including DC1 and DC2) and the expression of human leukocyte antigen DR (HLA DR) among PBC patients and controls<sup>98</sup>. NK cell-mediated innate immune responses may be crucial in the initial stages of PBC. An analysis of NK cells isolated from the spleen of PBC patients suggested that NK cells might activate autoreactive CD4<sup>+</sup> T cells by secreting IFN- $\gamma$  to destroy biliary epithelial cells<sup>96</sup>. CD4<sup>+</sup> T cells were enriched at IBD including CD. Certainly, as mentioned in the preceding text, CD4<sup>+</sup> T cells mediated the protective mechanism through the risk SNP. These results may provide evidence that the difference in immune activation may also contribute to explain the protective mechanism.

There are some limitations in our study. Firstly, to prevent population stratification, all

datasets are only from European population. Hence, there existed limited generalizability to these other populations. Secondly, although we revealed the potential shared genetic architecture but how the shared biological pathways work warrants further research. Thirdly, the sample size for PBC was relatively smaller, which might influence the results to some extent.

## **Conclusion**

In summary, we firstly confirmed the mutual protective correlation and negative genetic correlation between IBD and PBC. We identified novel shared risk SNPs and further explored the shared genetic overlap in tissue and cell-type levels between IBD and PBC. These findings may help to better explain the reason of rare comorbidity between IBD and PBC and may contribute to the targeted treatment and clinical management.

## **Data source**

GWAS summary statistics for PBC are available by application from:

<https://www.ebi.ac.uk/gwas/>.

GWAS summary statistics for PBC for replication are available by application from:

<https://gwas.mrcieu.ac.uk/>.

GWAS summary statistics for IBD are available by application from:

<https://www.ebi.ac.uk/gwas/>.

GEO dataset:

PBC: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi>

IBD: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE179285>

GTEX-seq:

[https://alkesgroup.broadinstitute.org/LDSCORE/LDSC\\_SEG\\_ldscores](https://alkesgroup.broadinstitute.org/LDSCORE/LDSC_SEG_ldscores)

scRNA-seq data:

Whole blood:

Lung and Spleen: <https://doi.org/10.1186/s13059-019-1906-x>;

Small Intestinal Epithelium:

[https://singlecell.broadinstitute.org/single\\_cell/study/SCP44/small-intestinal-epithelium](https://singlecell.broadinstitute.org/single_cell/study/SCP44/small-intestinal-epithelium)

### **Code availability**

LDSC: <https://github.com/bulik/ldsc>.

PLINK: <https://www.cog-genomics.org/plink/1.9>.

MTAG: <https://github.com/JonJala/mtag>.

CPASSOC: <http://hal.case.edu/~xxz10/zhuweb/>.

GSMR: <http://cnsgenomics.com/software/gsmr/>.

TwoSampleMR: <https://mrcieu.github.io/TwoSampleMR/>.

LDSC-SEG: <https://github.com/bulik/ldsc/wiki/Cell-type-specific-analyses>

MAGMA Celltyping: [https://neurogenomics.github.io/MAGMA\\_Celltyping](https://neurogenomics.github.io/MAGMA_Celltyping)

### **Author contributions**

WTH, RJ, RJZ and HC designed, revised and supervised the study; WTH, RJ, RJZ and XJW analyzed and organized the data; YYM, LJZ, SST and YLL generated the figures and tables; WTH, RJ, RJZ and HC wrote this manuscript; All authors reviewed and approved the final manuscript.

### **Funding**

This work was funded by the National Natural Science Foundation of China (82171698, 82170561, 81300279, and 81741067), the Natural Science Foundation for Distinguished Young Scholars of Guangdong Province (2021B1515020003), Natural Science Foundation of Guangdong Province (2022A1515012081), the Climbing Program of Introduced Talents and High-level Hospital Construction Project of Guangdong Provincial People's Hospital (DFJH201803, KJ012019099, KJ012021143, and KY012021183).

### **Acknowledgements**

We gratefully acknowledge the contributions from public available databases and the participants who contributed to those studies.



### Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### References

1. McGovern, D. P. B., Kugathasan, S. & Cho, J. H. Genetics of Inflammatory Bowel Diseases. *Gastroenterology* **149**, 1163-1176.e2 (2015).
2. Torres, J., Mehandru, S., Colombel, J.-F. & Peyrin-Biroulet, L. Crohn's disease. *Lancet* **389**, 1741–1755 (2017).
3. Peloquin, J. M., Goel, G., Villablanca, E. J. & Xavier, R. J. Mechanisms of Pediatric Inflammatory Bowel Disease. *Annu Rev Immunol* **34**, 31–64 (2016).
4. I, O., L, E., M, T., Dc, B. & Wj, S. Ulcerative colitis. *Lancet (London, England)* **380**, (2012).
5. Harbord, M. *et al.* The First European Evidence-based Consensus on Extra-intestinal Manifestations in Inflammatory Bowel Disease. *J Crohns Colitis* **10**, 239–254 (2016).
6. Xie, Y. *et al.* Causal Linkage Between Inflammatory Bowel Disease and Primary Sclerosing Cholangitis: A Two-Sample Mendelian Randomization Analysis. *Frontiers in Genetics* **12**, (2021).
7. Lleo, A., Leung, P. S. C., Hirschfield, G. M. & Gershwin, E. M. The Pathogenesis of Primary Biliary Cholangitis: A Comprehensive Review. *Semin Liver Dis* **40**, 34–48 (2020).
8. Kaplan, M. M. & Gershwin, M. E. Primary biliary cirrhosis. *N Engl J Med* **353**, 1261–1273 (2005).
9. Floreani, A. *et al.* Extrahepatic autoimmune conditions associated with primary biliary cirrhosis. *Clin Rev Allergy Immunol* **48**, 192–197 (2015).

10. Koulentaki, M. *et al.* Ulcerative colitis associated with primary biliary cirrhosis. *Dig Dis Sci* **44**, 1953–1956 (1999).
11. Liberal, R., Gaspar, R., Lopes, S. & Macedo, G. Primary biliary cholangitis in patients with inflammatory bowel disease. *Clin Res Hepatol Gastroenterol* **44**, e5–e9 (2020).
12. Efe, C. *et al.* Extrahepatic autoimmune diseases in primary biliary cholangitis: Prevalence and significance for clinical presentation and disease outcome. *J Gastroenterol Hepatol* **36**, 936–942 (2021).
13. Jang, H. J. *et al.* Development of primary biliary cirrhosis in a patient with Crohn’s disease: a case report and review of the literature. *Dig Dis Sci* **50**, 2335–2337 (2005).
14. Yoshida, E. M. & Erb, S. R. Primary biliary cirrhosis and inflammatory bowel disease. *Am J Gastroenterol* **88**, 1977–1978 (1993).
15. Sisman, G., Erzin, Y. & Bal, K. Primary biliary cirrhosis developing in a patient with Crohn’s disease during the course of infliximab treatment: the first case in the literature. *J Crohns Colitis* **7**, e397-398 (2013).
16. Invernizzi, P. & Gershwin, M. E. Primary biliary cirrhosis: bad genes, bad luck. *Dig Dis Sci* **57**, 599–601 (2012).
17. Rosen, M. J., Dhawan, A. & Saeed, S. A. Inflammatory Bowel Disease in Children and Adolescents. *JAMA Pediatr* **169**, 1053–1060 (2015).
18. Hodson, R. Inflammatory bowel disease. *Nature* **540**, S97 (2016).
19. Colombel, J.-F. Decade in review-IBD: IBD-genes, bacteria and new therapeutic strategies. *Nat Rev Gastroenterol Hepatol* **11**, 652–654 (2014).
20. Invernizzi, P. *et al.* Human leukocyte antigen polymorphisms in Italian primary biliary

- cirrhosis: a multicenter study of 664 patients and 1992 healthy controls. *Hepatology* **48**, 1906–1912 (2008).
21. Mells, G. F. *et al.* Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. *Nat Genet* **43**, 329–332 (2011).
  22. Liu, X. *et al.* Genome-wide meta-analyses identify three loci associated with primary biliary cirrhosis. *Nat Genet* **42**, 658–660 (2010).
  23. Hirschfield, G. M. *et al.* Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants. *N Engl J Med* **360**, 2544–2555 (2009).
  24. Colombel, J.-F. Decade in review-IBD: IBD-genes, bacteria and new therapeutic strategies. *Nat Rev Gastroenterol Hepatol* **11**, 652–654 (2014).
  25. Finucane, H. K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat Genet* **50**, 621–629 (2018).
  26. Yang, Y. *et al.* Investigating the shared genetic architecture between multiple sclerosis and inflammatory bowel diseases. *Nat Commun* **12**, 5641 (2021).
  27. Hashimoto, E. *et al.* Immunohistochemical characterization of hepatic lymphocytes in primary biliary cirrhosis in comparison with primary sclerosing cholangitis and autoimmune chronic active hepatitis. *Mayo Clin Proc* **68**, 1049–1055 (1993).
  28. Friedrich, M., Pohin, M. & Powrie, F. Cytokine Networks in the Pathophysiology of Inflammatory Bowel Disease. *Immunity* **50**, 992–1006 (2019).
  29. Zenewicz, L. A., Antov, A. & Flavell, R. A. CD4 T-cell differentiation and inflammatory bowel disease. *Trends in Molecular Medicine* **15**, 199–207 (2009).
  30. Cao, H. *et al.* The Pathogenicity and Synergistic Action of Th1 and Th17 Cells in

- Inflammatory Bowel Diseases. *Inflamm Bowel Dis* **izac199** (2022) doi:10.1093/ibd/izac199.
31. Giuffrida, P. & Di Sabatino, A. Targeting T cells in inflammatory bowel disease. *Pharmacol Res* **159**, 105040 (2020).
  32. Tindemans, I., Joosse, M. E. & Samsom, J. N. Dissecting the Heterogeneity in T-Cell Mediated Inflammation in IBD. *Cells* **9**, 110 (2020).
  33. de Lange, K. M. *et al.* Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* **49**, 256–261 (2017).
  34. Cordell, H. J. *et al.* An international genome-wide meta-analysis of primary biliary cholangitis: Novel risk loci and candidate drugs. *J Hepatol* **75**, 572–581 (2021).
  35. Cordell, H. J. *et al.* International genome-wide meta-analysis identifies new primary biliary cirrhosis risk loci and targetable pathogenic pathways. *Nat Commun* **6**, 8019 (2015).
  36. Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum Mol Genet* **27**, 3641–3649 (2018).
  37. GTEx Consortium *et al.* Genetic effects on gene expression across human tissues. *Nature* **550**, 204–213 (2017).
  38. Keir, M. E. *et al.* Regulation and Role of  $\alpha$ E Integrin and Gut Homing Integrins in Migration and Retention of Intestinal Lymphocytes during Inflammatory Bowel Disease. *J Immunol* **207**, 2245–2254 (2021).
  39. Nakagawa, R. *et al.* miR-425 regulates inflammatory cytokine production in CD4+ T cells via N-Ras upregulation in primary biliary cholangitis. *J Hepatol* **66**, 1223–1230 (2017).
  40. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat Commun* **8**, 14049 (2017).

41. Haber, A. L. *et al.* A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333–339 (2017).
42. Xie, X. *et al.* Single-cell transcriptomic landscape of human blood cells. *Natl Sci Rev* **8**, nwaa180 (2021).
43. Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat Commun* **9**, 224 (2018).
44. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol* **40**, 304–314 (2016).
45. Burgess, S. & Thompson, S. G. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol* **32**, 377–389 (2017).
46. Hartwig, F. P., Davey Smith, G. & Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* **46**, 1985–1998 (2017).
47. Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* **37**, 658–665 (2013).
48. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **7**, e34408 (2018).
49. Verbanck, M., Chen, C.-Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet* **50**, 693–698 (2018).
50. Liu, G. *et al.* PICALM gene rs3851179 polymorphism contributes to Alzheimer’s disease in

- an Asian population. *Neuromolecular Med* **15**, 384–388 (2013).
51. Greco M, F. D., Minelli, C., Sheehan, N. A. & Thompson, J. R. Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome. *Stat Med* **34**, 2926–2940 (2015).
  52. Wang, R. Mendelian randomization study updates the effect of 25-hydroxyvitamin D levels on the risk of multiple sclerosis. *J Transl Med* **20**, 3 (2022).
  53. Noyce, A. J. *et al.* Estimating the causal influence of body mass index on risk of Parkinson disease: A Mendelian randomisation study. *PLoS Med* **14**, e1002314 (2017).
  54. Hk, F. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature genetics* **47**, (2015).
  55. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
  56. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236–1241 (2015).
  57. Yao, Y. *et al.* An atlas of genetic correlations between gestational age and common psychiatric disorders. *Autism Res* **15**, 1008–1017 (2022).
  58. Lu, Q. *et al.* A Powerful Approach to Estimating Annotation-Stratified Genetic Covariance via GWAS Summary Statistics. *Am J Hum Genet* **101**, 939–964 (2017).
  59. H, S., N, M., S, S. & B, P. Local Genetic Correlation Gives Insights into the Shared Genetic Architecture of Complex Traits. *American journal of human genetics* **101**, (2017).
  60. Berisa, T. & Pickrell, J. K. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics* **32**, 283–285 (2016).

61. Stolp Andersen, M. *et al.* Dissecting the limited genetic overlap of Parkinson's and Alzheimer's disease. *Ann Clin Transl Neurol* **9**, 1289–1295 (2022).
62. Zhu, X. *et al.* Meta-analysis of correlated traits via summary statistics from GWASs with an application in hypertension. *Am J Hum Genet* **96**, 21–36 (2015).
63. 23andMe Research Team *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat Genet* **50**, 229–237 (2018).
64. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* **11**, e1004219 (2015).
65. Skene, N. G. & Grant, S. G. N. Identification of Vulnerable Cell Types in Major Brain Disorders Using Single Cell Transcriptomes and Expression Weighted Cell Type Enrichment. *Front Neurosci* **10**, 16 (2016).
66. Gazal, S. *et al.* Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nat Genet* **49**, 1421–1427 (2017).
67. Xiao, W.-B. & Liu, Y.-L. Primary biliary cirrhosis and ulcerative colitis: A case report and review of literature. *World J Gastroenterol* **9**, 878–880 (2003).
68. Arai, O., Ikeda, H., Mouri, H., Notohara, K. & Matsueda, K. Two cases of inflammatory bowel disease diagnosed in the course of primary biliary cirrhosis. *Nihon Shokakibyō Gakkai Zasshi* **107**, 900–908 (2010).
69. Polychronopoulou, E., Lygoura, V., Gatselis, N. K. & Dalekos, G. N. Increased cholestatic enzymes in two patients with long-term history of ulcerative colitis: consider primary biliary cholangitis not always primary sclerosing cholangitis. *BMJ Case Rep* **2017**, bcr2017220824 (2017).

70. Fousekis, F. S., Theopistos, V. I., Katsanos, K. H., Tsianos, E. V. & Christodoulou, D. K. Hepatobiliary Manifestations and Complications in Inflammatory Bowel Disease: A Review. *Gastroenterology Res* **11**, 83–94 (2018).
71. Lleo, A. *et al.* Y chromosome loss in male patients with primary biliary cirrhosis. *Journal of Autoimmunity* **41**, 87–91 (2013).
72. Lleo, A. *et al.* DNA methylation profiling of the X chromosome reveals an aberrant demethylation on CXCR3 promoter in primary biliary cirrhosis. *Clin Epigenet* **7**, 61 (2015).
73. Hitomi, Y. *et al.* rs1944919 on chromosome 11q23.1 and its effector genes COLCA1/COLCA2 confer susceptibility to primary biliary cholangitis. *Sci Rep* **11**, 4557 (2021).
74. Hitomi, Y. *et al.* POGUT1, the putative effector gene driven by rs2293370 in primary biliary cholangitis susceptibility locus chromosome 3q13.33. *Sci Rep* **9**, 102 (2019).
75. H, K. *et al.* Quantitative and functional analysis of PDC-E2-specific autoreactive cytotoxic T lymphocytes in primary biliary cirrhosis. *The Journal of clinical investigation* **109**, (2002).
76. La, Z., A, A. & Ra, F. CD4 T-cell differentiation and inflammatory bowel disease. *Trends in molecular medicine* **15**, (2009).
77. Ramon, H. E. *et al.* The ubiquitin ligase adaptor Ndfip1 regulates T cell-mediated gastrointestinal inflammation and inflammatory bowel disease susceptibility. *Mucosal Immunol* **4**, 314–324 (2011).
78. Altin, J. A. *et al.* Ndfip1 mediates peripheral tolerance to self and exogenous antigen by inducing cell cycle exit in responding CD4+ T cells. *Proc Natl Acad Sci U S A* **111**, 2067–2074 (2014).



79. Söderman, J., Berglind, L. & Almer, S. Gene Expression-Genotype Analysis Implicates GSDMA, GSDMB, and LRRC3C as Contributors to Inflammatory Bowel Disease Susceptibility. *Biomed Res Int* **2015**, 834805 (2015).
80. Franke, A. *et al.* Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* **42**, 1118–1125 (2010).
81. Fazio, A. *et al.* DNA methyltransferase 3A controls intestinal epithelial barrier function and regeneration in the colon. *Nat Commun* **13**, 6266 (2022).
82. Spalinger, M. R., McCole, D. F., Rogler, G. & Scharl, M. Role of Protein Tyrosine Phosphatases in Regulating the Immune System: Implications for Chronic Intestinal Inflammation. *Inflammatory Bowel Diseases* **21**, 645–655 (2015).
83. Li, K., Gesang, L., Dan, Z. & Gusang, L. Genome-Wide Transcriptional Analysis Reveals the Protection against Hypoxia-Induced Oxidative Injury in the Intestine of Tibetans via the Inhibition of GRB2/EGFR/PTPN11 Pathways. *Oxidative Medicine and Cellular Longevity* **2016**, e6967396 (2016).
84. Hitomi, Y. *et al.* Identification of the functional variant driving ORMDL3 and GSDMB expression in human chromosome 17q12-21 in primary biliary cholangitis. *Sci Rep* **7**, 2904 (2017).
85. Han, C. *et al.* FoxO1 regulates TLR4/MyD88/MD2-NF- $\kappa$ B inflammatory signalling in mucosal barrier injury of inflammatory bowel disease. *Journal of Cellular and Molecular Medicine* **24**, 3712–3723 (2020).
86. Yang, X. *et al.* MicroRNA-425 facilitates pathogenic Th17 cell differentiation by targeting forkhead box O1 (Foxo1) and is associated with inflammatory bowel disease. *Biochem*

- Biophys Res Commun* **496**, 352–358 (2018).
87. Shadrin, A. A. *et al.* Shared Genetics of Multiple System Atrophy and Inflammatory Bowel Disease. *Mov Disord* **36**, 449–459 (2021).
88. Liu, J. Z. *et al.* Dense fine-mapping study identifies new susceptibility loci for primary biliary cirrhosis. *Nat Genet* **44**, 1137–1141 (2012).
89. Jones, D. E. J. Pathogenesis of primary biliary cirrhosis. *J Hepatol* **39**, 639–648 (2003).
90. Mateer, S. W. *et al.* Potential mechanisms regulating pulmonary pathology in inflammatory bowel disease. *J Leukoc Biol* **98**, 727–737 (2015).
91. Idilman, I. S. *et al.* Magnetic resonance imaging features in 283 patients with primary biliary cholangitis. *Eur Radiol* **30**, 5139–5148 (2020).
92. Wen, Z. & Fiocchi, C. Inflammatory bowel disease: autoimmune or immune-mediated pathogenesis? *Clin Dev Immunol* **11**, 195–204 (2004).
93. Steinbach, E. C. & Plevy, S. E. The role of macrophages and dendritic cells in the initiation of inflammation in IBD. *Inflamm Bowel Dis* **20**, 166–175 (2014).
94. Casalegno Garduño, R. & Däbritz, J. New Insights on CD8+ T Cells in Inflammatory Bowel Disease and Therapeutic Approaches. *Front Immunol* **12**, 738762 (2021).
95. Noble, A. *et al.* Deficient Resident Memory T Cell and CD8 T Cell Response to Commensals in Inflammatory Bowel Disease. *Journal of Crohn's and Colitis* **14**, 525–537 (2020).
96. Shimoda, S. *et al.* Natural killer cells regulate T cell immune responses in primary biliary cirrhosis. *Hepatology* **62**, 1817–1827 (2015).
97. Dhirapong, A. *et al.* Therapeutic effect of cytotoxic T lymphocyte antigen 4/immunoglobulin on a murine model of primary biliary cirrhosis. *Hepatology* **57**, 708–715 (2013).

98. Dendritic cell subtypes in autoimmune liver diseases; decreased expression of HLA DR and CD123 on type 2 dendritic cells. *Hepatology Research* **22**, 241–249 (2002).

### Figure legend

**Figure. 1 The estimation of causality between PBC and IBD (Including CD and UC) based on five Mendelian Randomization analysis methods.** Five boxes represent GSMR: Generalized Summary-data-based Mendelian Randomization, MR Egger, IVW: Inverse variance weighting, Weighted Median and Weighted Mode from the top to bottom. IBD: inflammatory bowel disease, CD: Crohn's disease, UC: ulcerative colitis, PBC: primary biliary cholangitis.

**Figure. 2 Local genetic correlation at loci ascertained for GWAS hits ( $p < 5 \times 10^{-6}$ ) specific to PBC and IBD(A), CD (B) and UC (right), respectively.** IBD: inflammatory bowel disease, CD: Crohn's disease, UC: ulcerative colitis, PBC: primary biliary cholangitis.

**Figure. 3 Tissue type-specific enrichment SNP heritability for (A) inflammatory bowel disease, (B) Crohn's disease, (C) ulcerative colitis and (D) primary biliary cholangitis.** The x axis represents negative  $\log_{10}$  FDR-adjusted  $P$ -values for each test, with the grey dotted lines indicating a 5% FDR. FDR: false discovery rate, SNP: single nucleotide polymorphism.

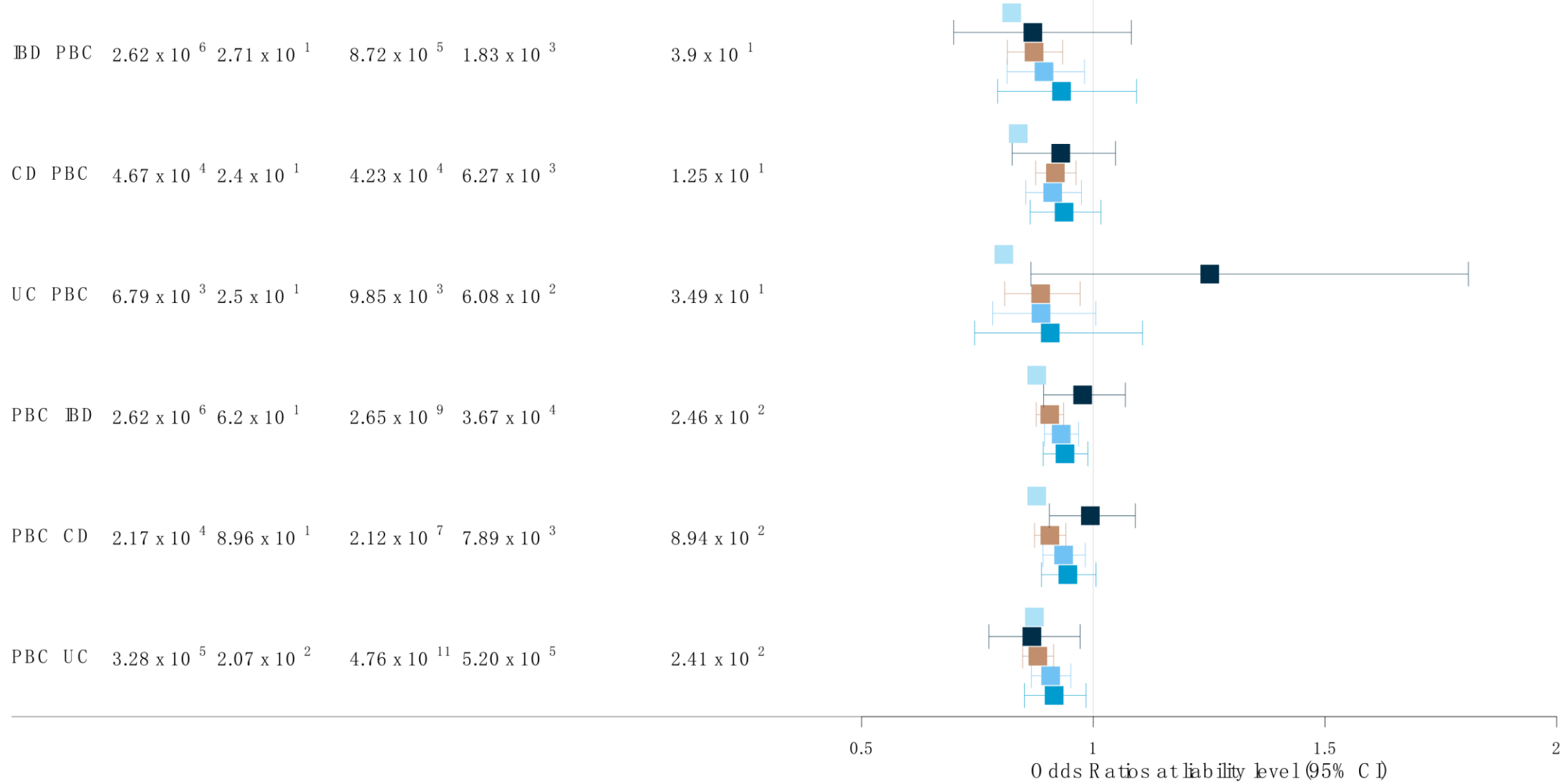
**Figure. 4 Cell type-specific enrichment for (A) inflammatory bowel disease, (B) Crohn's disease and (C) primary biliary cholangitis, analyzed by MAGMA method.** The x axis represents negative  $\log_{10}$  FDR-adjusted  $P$ -values for each test, with the grey dotted lines indicating a 5% FDR. FDR: false discovery rate, SNP: single nucleotide polymorphism, MAGMA: multi-marker analysis of GenoMic annotation.

**Table 1. Heritability and genetic correlation estimated for PBC and IBD including CD and UC using LDSC and GNOVA.** PBC: primary biliary cholangitis, IBD: inflammatory bowel disease, UC: ulcerative colitis, CD: Crohn's disease, LDSC: LD score regression, GNOVA: Genetic covariance analyzer.

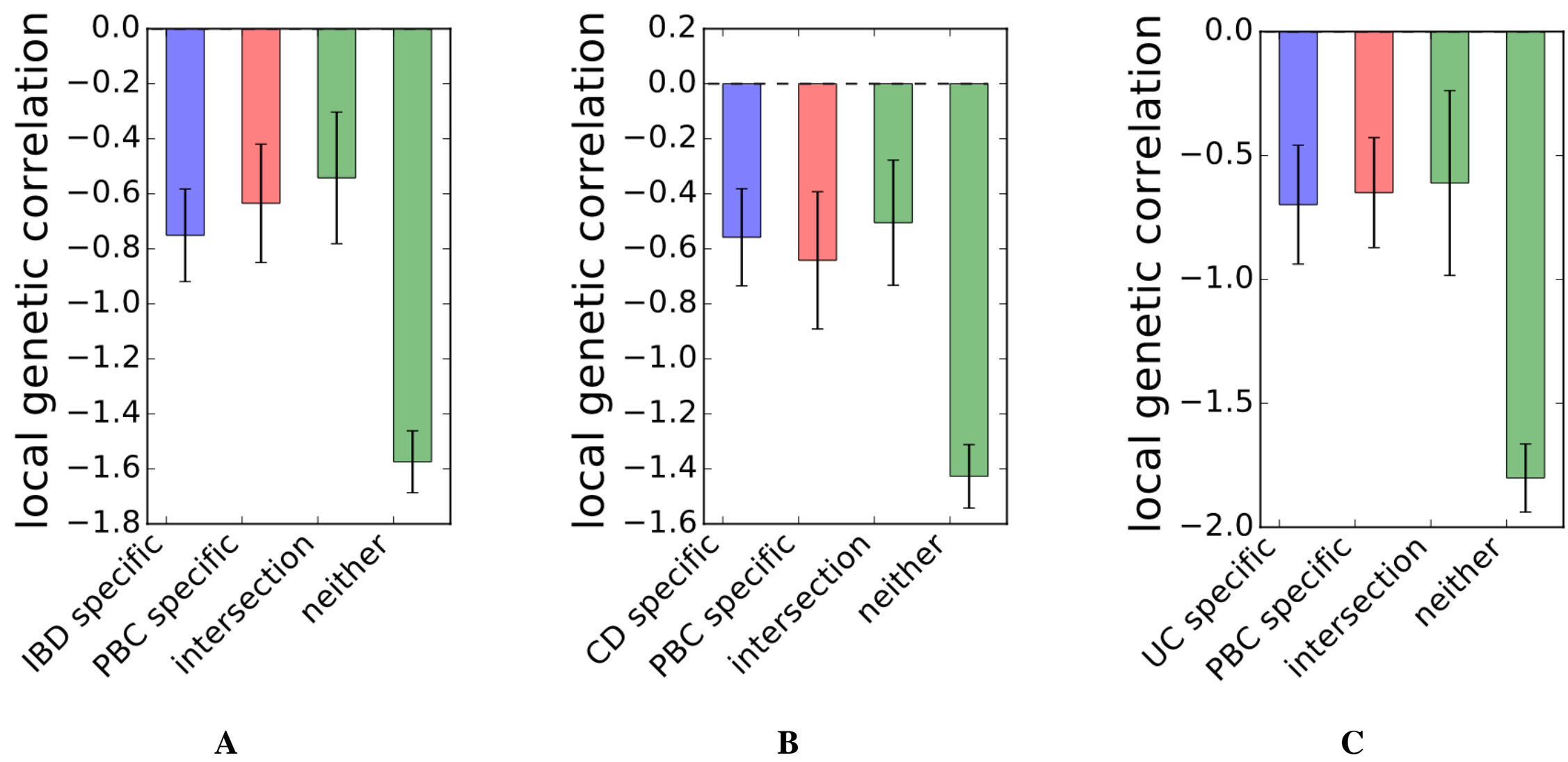
**Table 2. Shared risk SNPs associated with cross-trait PBC and IBD (or UC or CD) presented by MTAG, CPASSOC and  $\rho$ -HESS.** PBC: primary biliary cholangitis, IBD: inflammatory bowel disease, UC: ulcerative colitis, CD: Crohn's disease, SNP: single nucleotide polymorphisms, BP: base pair, MTAG: Multi-Trait Analysis of genome-wide association study (GWAS), CPASSOC: Cross Phenotype Association, HESS: Heritability Estimation from Summary Statistics, CHR: chromosome.

# Figure 1

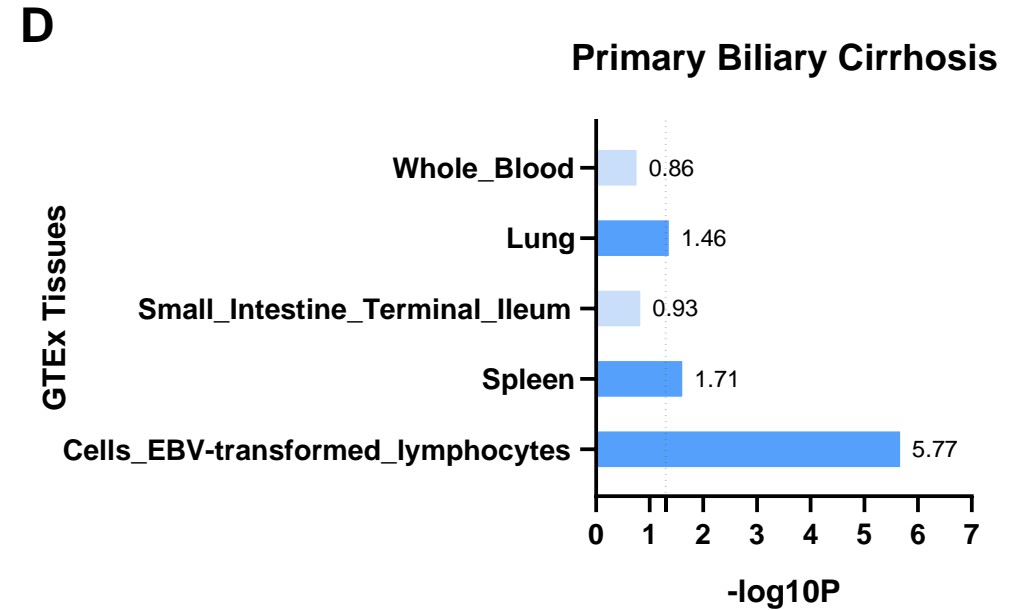
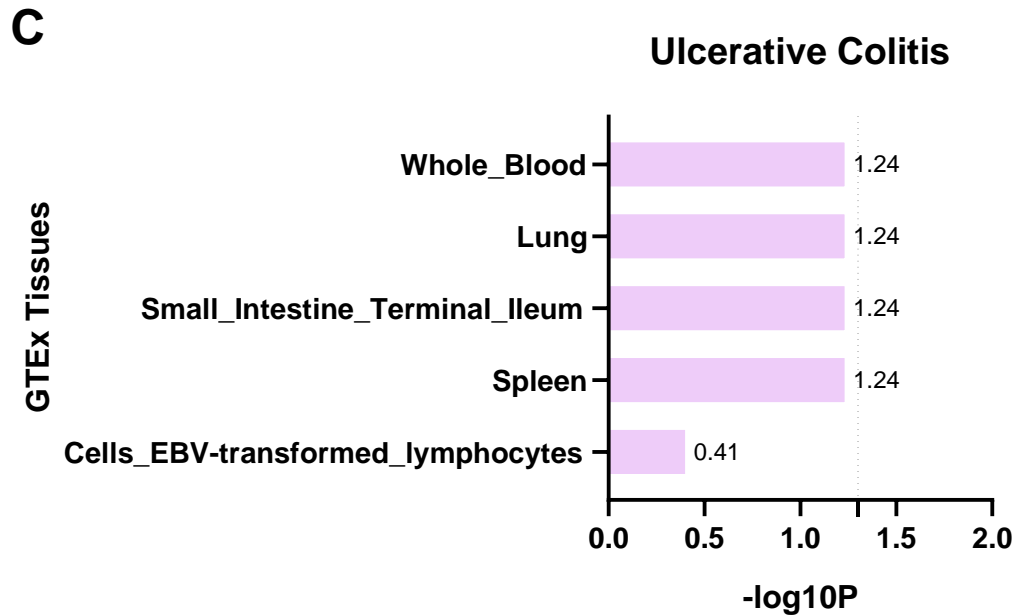
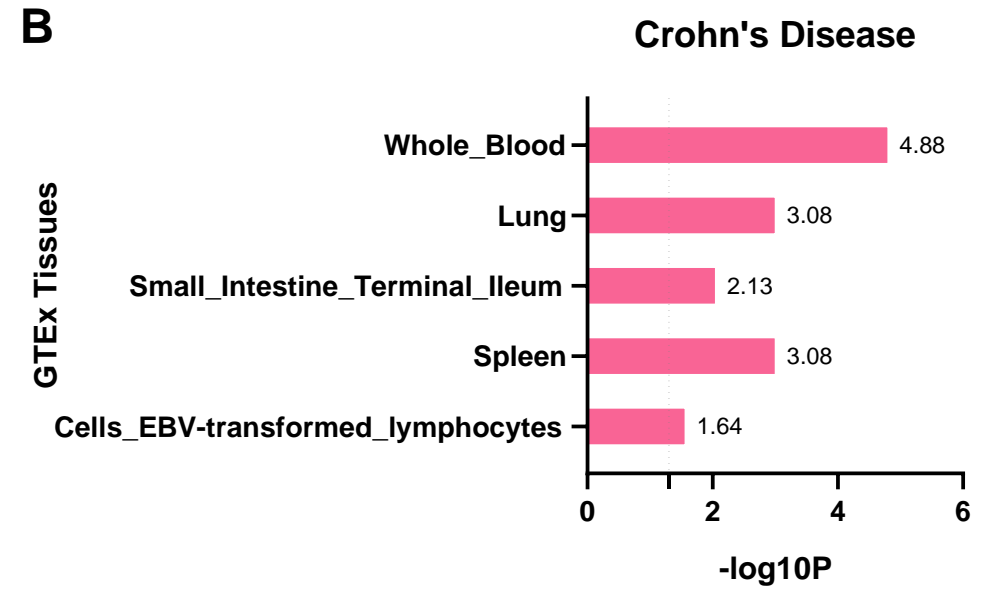
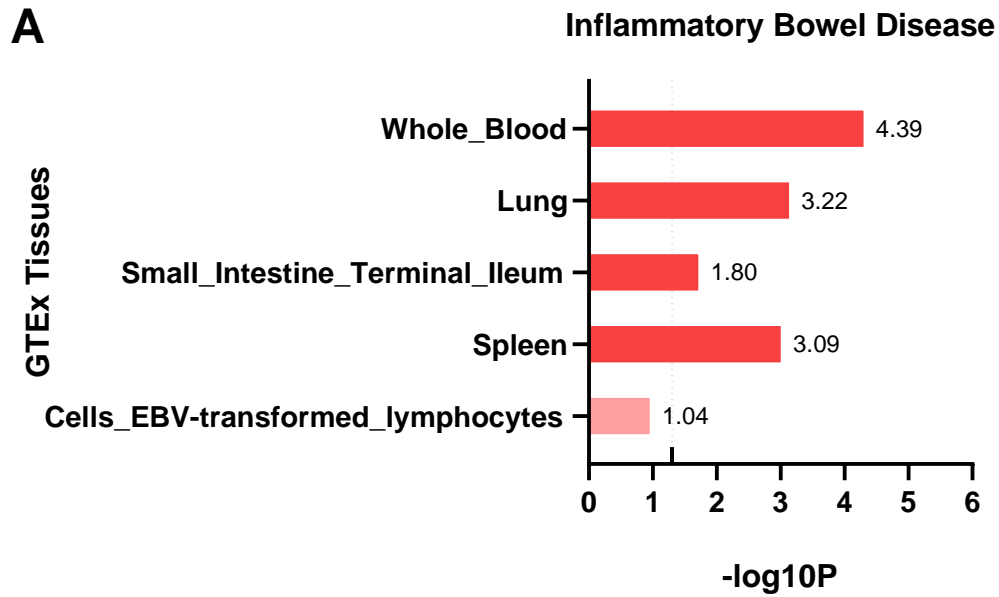
Exp Out GSMR P MR Egger P IVW P Weighted Median P Weighted Mode P



**Figure 2**



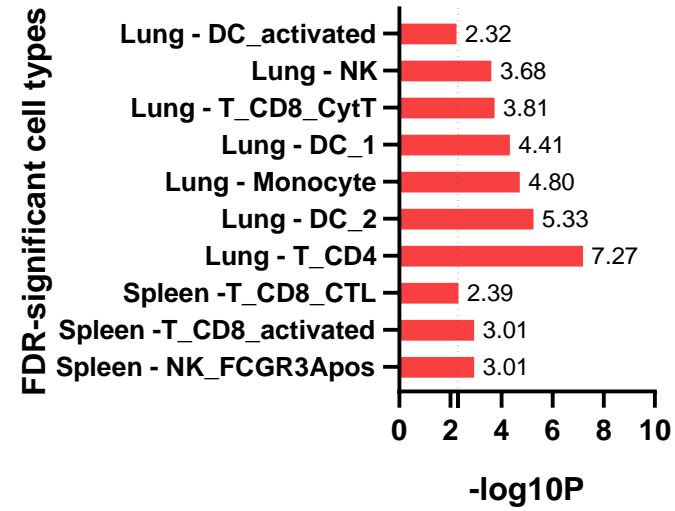
# Figure 3



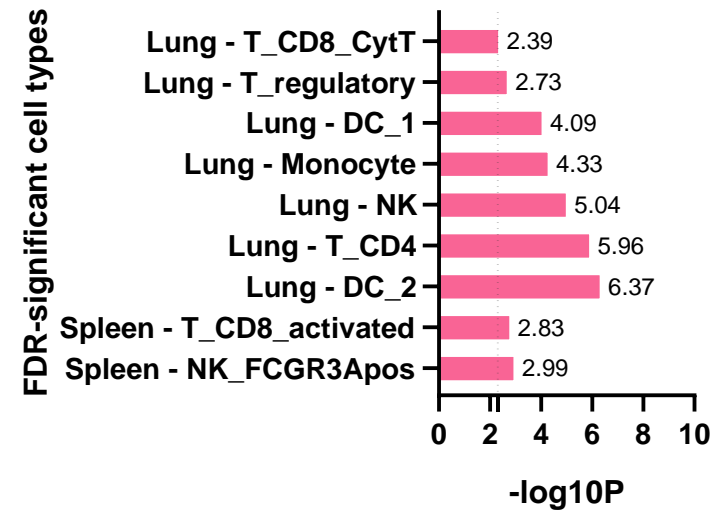
# Figure 4

**A**

## Inflammatory Bowel Disease

**B**

## Crohn's Disease

**C**

## Primary Biliary Cirrhosis

