

# SCIENTIFIC REPORTS

OPEN

## Common functional alterations identified in blood transcriptome of autoimmune cholestatic liver and inflammatory bowel diseases

Jerzy Ostrowski<sup>1,2</sup>, Krzysztof Goryca<sup>1</sup>, Izabella Lazowska<sup>3</sup>, Agnieszka Rogowska<sup>2</sup>, Agnieszka Paziewska<sup>2</sup>, Michalina Dabrowska<sup>1</sup>, Filip Ambrozkiewicz<sup>1</sup>, Jakub Karczmariski<sup>1</sup>, Aneta Balabas<sup>1</sup>, Anna Kluska<sup>1</sup>, Magdalena Piatkowska<sup>1</sup>, Natalia Zeber-Lubecka<sup>2</sup>, Maria Kulecka<sup>2</sup>, Andrzej Habior<sup>2</sup>, Michal Mikula<sup>1</sup>, The Polish PBC study Group\* & The Polish IBD study Group\*

Primary biliary cholangitis (PBC), primary sclerosing cholangitis (PSC), and inflammatory bowel diseases (IBDs), including Crohn's disease (CD) and ulcerative colitis (UC), are heterogeneous chronic autoimmune diseases that may share underlying pathogenic mechanisms. Herein, we compared simultaneously analyzed blood transcriptomes from patients with PBC, PSC, and IBD. Microarray-based measurements were conducted using RNA isolated from whole blood samples from 90, 45, 95 and 93 patients with PBC, PSC, CD, and UC, respectively, and 47 healthy controls. Expression levels of selected transcripts were analyzed by quantitative reverse-transcribed PCR using an independent cohort of 292, 71 and 727 patients with PBC, PSC, and IBD, respectively. Of 4026, 2650 and 4967 probe sets differentially expressed (adjusted p-value < 0.05) in samples from patients with PBC, PSC, and IBD, respectively, compared with healthy controls, 1946 were common to all three comparisons. Functional analyses indicated that most terms enriched for genes differentially expressed in PBC, PSC, and IBD patients compared with healthy controls were related to mitochondrial function, the vesicle endomembrane system, and GTPase-mediated processes. This study indicates that microarray-based profiling of blood gene expression supports research into the molecular mechanisms underlying disease, rather than being useful for selection of diagnostic biomarkers for use in clinical practice.

Primary biliary cholangitis (PBC), primary sclerosing cholangitis (PSC), and inflammatory bowel diseases (IBDs), including Crohn's disease (CD) and ulcerative colitis (UC), are heterogeneous chronic autoimmune diseases with genetic, immunologic, and environmental components. Genetic risk factors for these conditions are primarily non-protein-coding single nucleotide polymorphisms with similar small effect sizes<sup>1-3</sup>.

PBC is characterized by lymphoplasmacytic infiltration around the interlobular ducts of the liver, resulting in progressive immune-mediated destruction of interlobular biliary ductules associated with a classical feature of autoimmune conditions, antimitochondrial antibodies. PSC manifests as cholangiocytic injuries associated with nonspecific inflammation. In both of these cholangiopathies, progressive fibrous obliteration of the intrahepatic and extrahepatic biliary tree results in chronic cholestasis leading to liver cirrhosis<sup>4-6</sup>. PBC occurs more frequently in women than men and primarily in middle age, with prevalence rates ranging from 40 to 400 patients per million and an incidence range of 0.7 to 49 per million<sup>7-10</sup>. PSC affects 9 to 13 patients per million annually with a male-to-female ratio of 2:1<sup>11</sup>. Up to 80% of PSC cases are associated with IBD, while PSC is present in 3–8% of all patients with UC and 1–3% of patients with CD<sup>12,13</sup>.

<sup>1</sup>Department of Genetics, Maria Skłodowska-Curie Institute – Oncology Centre, Warsaw, 02-781, Poland.

<sup>2</sup>Department of Gastroenterology and Hepatology, Medical Center for Postgraduate Education, Warsaw, 02-781, Poland. <sup>3</sup>Department of Pediatric Gastroenterology and Nutrition, Medical University of Warsaw, Warsaw, 02-091, Poland. \*A comprehensive list of consortium members appears at the end of the paper. Correspondence and requests for materials should be addressed to J.O. (email: [jostrow@warman.com.pl](mailto:jostrow@warman.com.pl)) or A.H. (email: [habior.andrzej@gmail.com](mailto:habior.andrzej@gmail.com))

Correspondence and requests for materials should be addressed to J.O. (email: [jostrow@warman.com.pl](mailto:jostrow@warman.com.pl)) or A.H. (email: [habior.andrzej@gmail.com](mailto:habior.andrzej@gmail.com))

|                     |               | PBC   | PSC   | IBDs     |        |        |       | Controls |
|---------------------|---------------|-------|-------|----------|--------|--------|-------|----------|
|                     |               |       |       | Children |        | Adults |       |          |
|                     |               |       |       | CD       | UC     | CD     | UC    |          |
| Discovery cohorts   | Females/Males | 90/0  | 19/27 | 29/21    | 28/19  | 23/22  | 26/20 | 26/21    |
|                     | Age; range    | 29–76 | 16–53 | 2–17     | 1–17   | 19–69  | 21–66 | 38–62    |
|                     | Age; median   | 57    | 29    | 13       | 15     | 34     | 36    | 46       |
| Replication cohorts | Females/Males | 292/0 | 14/56 | 103/150  | 111/88 | 86/54  | 80/55 | 107/99   |
|                     | Age; range    | 18–85 | 17–64 | 2–17     | 1–17   | 18–70  | 18–73 | 29–78    |
|                     | Age; median   | 56    | 31    | 15       | 15     | 29     | 35    | 49       |

**Table 1.** Summary of the main epidemiological variables for the discovery and replication cohorts.

IBDs result from multiple intestinal immunopathological processes, in which Th17 cells have a central role, in response to host intestinal microflora that induce the initiation and maintenance of intestinal inflammation<sup>14</sup>. Of the two major types of IBD, UC is characterized by inflammation extending continuously from the rectum along the entire colon, while in CD the inflammatory response is typically localized to the distal small intestine and colon. In UC, inflammation is confined to the mucosal surface of the colon, while in CD it is transmural. IBD onset can occur from early childhood to beyond the sixth decade of life, with childhood-onset IBD representing 10–25% of all cases<sup>15</sup>. Moreover, while PBC and PSC are progressive disorders, IBDs typically present as repeated cycles of relapse and remission of intestinal inflammation.

Blood comes into contact with the cells, tissues, and organs of the entire organism and constitutes a primary aspect of the immune defense system. Hence, it is not surprising that gene expression changes in white blood cells (WBCs) are associated with a wide range of pathological conditions. Blood can be considered as a surrogate for traditional tissue specimens employed for clinical diagnosis, and analyses of WBC expression profiles provide a non-invasive method that can be used to support investigations of both the molecular mechanisms underlying disease and medical practice<sup>16</sup>. Although the fibrous cholangiopathies, PBC and PSC, and nonspecific IBDs, UC and CD, exhibit significant differences in their clinical presentation, chronic inflammation and dysregulated immune responses are common to both types of disorder. Consequently, similar risk factors may be implicated in their pathogenesis, particularly given the crosstalk between bile acids (BAs) and gut microbes<sup>17</sup>. Once the bile duct and intestinal defense systems become affected, inappropriate innate immune and inflammatory responses may contribute to disturbed antibacterial reactive oxygen species (ROS)-mediated and mitochondrial autophagy. Whether alterations similar to those in tissues directly affected by disease can be observed within WBCs remains open to question. While several previous studies uncovered alterations of WBC gene expression in IBDs<sup>18–21</sup>, no comparable investigations of patients with PBC or PSC have been reported to date.

The main aim of our study was to uncover possible pathomechanisms common for PBC, PSC, UC, and CD by analysis of blood-based transcriptomes simultaneously generated for all of them. Additionally, basing on microarray profiling we intended to identify their new biomarkers. However, although our study selected aberrations of cellular signaling and regulatory pathways shared across all of the studied disorders, we did not select genes which could be used in a diagnostic screening.

## Materials and Methods

**Ethics approval and consent to participate.** The study was approved by the ethics committee (decision 46/PW/2011) of the Medical Center for Postgraduate Education, Warsaw, Poland. Informed consent was obtained from all subjects or, if subjects were under 18, from a parent and/or legal guardian. The study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki.

**Study subjects.** Patient cohorts included in this study comprised the following: 382 female patients with PBC, 331 of whom were antimitochondrial antibody positive; 116 (33 female and 83 male) patients with PSC; and 915 (486 females and 429 males) patients with IBD. Of patients with IBD, 488 (303 children) and 427 (246 children) were diagnosed with CD and UC, respectively. All enrolled patients and controls were Polish Caucasians. Most PSC patients were diagnosed with IBD: 12 with CD and 93 with UC. Diagnosis of PBC was based on standard clinical, biochemical, serological, and histological criteria, and PSC was diagnosed according to standard clinical, biochemical, cholangiographic, and (in some patients) histological criteria, according to the European Association for the Study of Liver (EASL)<sup>14</sup>. Before inclusion, all PBC patients and some PSC patients were treated with ursodeoxycholic acid; 20 PSC patients then underwent liver transplantation. IBDs were diagnosed using the Porto criteria, modified in accordance with the recommendations of the European Crohn's and Colitis Organization (ECCO) for children, and according to ECCO guidelines for adults. The CD activity index (CDAI), the UC activity index (UCAI), and their pediatric versions (PCDAI/PUCAI) were determined to evaluate disease severity<sup>22–24</sup>. Before inclusion most IBD patients were given mesalazine, and in majority of them the blood samples were collected before additional medication regimes (immunosuppressants, glucocorticoids, biologic therapy) were ordered. Blood samples from 184/82/253 healthy individuals served as controls for PBC/PSC/IBD, respectively. Mean AST/ALT for PBC patients was 3.4/3.1, respectively. Summaries of the main epidemiological variables for each group are presented in Table 1.

**RNA extraction.** For RNA extraction, whole blood was collected and total RNA was isolated using the Tempus RNA Isolation Kit (Thermo Fisher Scientific), according to the manufacturer's instructions. RNA quality

and quantity were analyzed using a NanoDrop spectrophotometer, and samples with A260/A280 ratios of 1.8–2.1 were further assessed using an Agilent 2100 Bioanalyzer. Samples used for microarray analysis had RNA integrity numbers in the range 7.6–9.6.

**Gene expression microarray analysis.** Whole-transcriptome profiling was performed by AROS Applied Biotechnology services, using an HT-12 v4 Expression BeadChip (Illumina, San Diego, CA, USA). The average bead signals from the chip were quantile normalized, with no background correction. All computations were performed using R 3.4.1 software with the Bioconductor extension<sup>25</sup>. Principal component analysis (PCA) was used for the initial quality inspection. 9 samples (3 PSC, 3 CU, 1 PBC, 1 CD and 1 control) were removed as outliers. Probe sets with expression detected (detection  $p$ -value  $< 0.05$ ) in less than 5 samples were discarded. The remaining measurements were filtered according to the ratio of the range between the 10th and 90th percentile (IQR10) and the median normalized IQR10 (NIQR10). Only probes with NIQR10 values higher than the median NIQR10 for the whole set were selected for analysis. Genes showing differential expression were selected according to  $p$ -value determined by  $t$ -test (Welch's variant) after correction for multiple hypothesis testing using the Benjamini–Hochberg algorithm. Adjusted  $p$ -values  $< 0.05$  were considered significant.

R code used for data analysis has been provided as Supplementary File 1.

**Quantitative reverse-transcribed PCR (qRT-PCR).** Quantitative reverse-transcribed PCR (qRT-PCR) was performed as described previously<sup>26</sup> using predesigned TaqMan Gene Expression assays or Sybr Green chemistry (Thermo Fisher Scientific). The geometric mean expression levels of *RPLP0* and *UBC* mRNAs were used as normalization factors. Gene expression levels were calculated using the  $\Delta\Delta C_t$  method<sup>27</sup>. Results were analyzed using the Mann–Whitney U-test in GraphPad Prism 5 software (GraphPad Software Inc., La Jolla, CA, USA), and  $p$ -values  $< 0.05$  were considered significant. The list of Taqman assays and primers is provided in Supplementary Table S1.

**Functional analysis.** Functional analyses were conducted in R (version 3.4.1). Gene Set Enrichment Analysis (GSEA) implemented in the *gseGO* function from *clusterProfiler* package (version 3.4.4)<sup>28</sup> was used to link gene expression profiles with Gene Ontology (GO) terms. GO terms were limited to those with between 100 and 300 genes mapped. *enrichPathway* function from *ReactomePA* package (version 1.20.2) was used to associate selected gene set with Reactome pathways. Resulting  $p$ -values were adjusted for multiple hypothesis testing using the Benjamini–Hochberg algorithm.

## Results

Transcriptome analysis was carried out using samples from patients with two cholestatic liver diseases (ChLDs), PBC and PSC, and two IBDs, CD and UC. While all of these disorders present unique clinicopathological features, ChLDs and IBDs may share underlying processes common to their pathogenesis. Microarray-based assays were conducted by hybridization of 370 RNA samples to Human HT-12 v4 Expression BeadChip microarrays. Of the 370 samples, 90, 45, 95, and 93 were from patients with PBC, PSC, CD, and UC, respectively, while 47 were from healthy controls. Transformation of gene expression variables from each array to their corresponding principal-component scores revealed that the consistency of the microarray data sets was as expected (Supplementary Fig. 1).

The number of differentially expressed genes detected for comparisons of CD and UC with controls were similar (Supplementary Table S2, 4649/4071, respectively). The concordance between the most significantly differentiating genes for both diseases was almost perfect with Spearman correlation coefficient equal 0.93 (Supplementary Fig. 2A) and higher than the correlation between PBC and PSC (Supplementary Fig. 2B). Therefore, while looking for the common functional alterations CD and UC were merged into single IBD group.

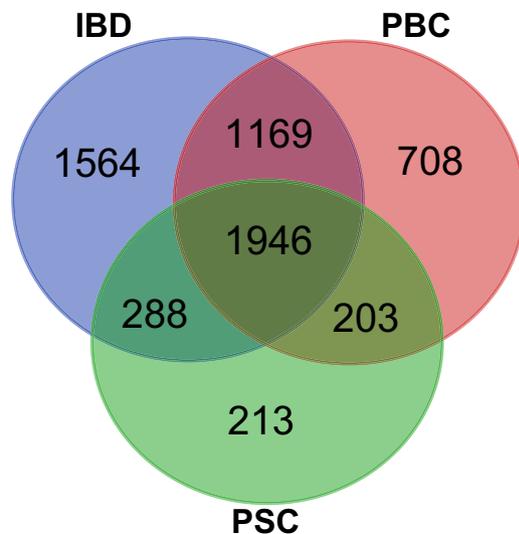
Although it is believed that the etiology of early and late onset IBD is different, the whole transcriptome expression pattern haven't differentiate children and adult patients (Supplementary Fig. 1C,D). Also, the most significant expression differences between each age group and controls were similar (Supplementary Fig. 2C,D).

According to pair-wise comparisons, 4026, 2650 and 4967 genes were differentially expressed between healthy controls and patients with PBC, PSC, and IBD (combined results of CD and UC), respectively (Fig. 1). Of these, 1946 genes were common to all three comparisons.

**Functional analysis according to GO subcategories.** Forty-three GO terms were over-represented among these common probe sets, 23, 12, and 8 of which were attributed to “biological process” (BP), “molecular function” (MF), and “cellular component” (CC) GO terms, respectively (Table 2). The majority of over-represented terms were related to mitochondrial respiration and ATP synthesis, with a few associated with signal transduction by small GTPases and membrane biogenesis and trafficking.

When the 1946 genes commonly dysregulated in all three disorders were annotated according to the Reactome signaling pathway database, 42 pathways were identified (Supplementary Table 3). Among these, the following terms exhibited the highest level of significance: R-HSA-163200, Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins (adjusted  $p = 3.01E-17$ ); R-HSA-1428517, The citric acid (TCA) cycle and respiratory electron transport (adjusted  $p = 1.59E-14$ ); R-HSA-611105, Respiratory electron transport (adjusted  $p = 4.18E-14$ ); R-HSA-6799198, Complex I biogenesis (adjusted  $p = 1.17E-09$ ); and R-HSA-5368286, Mitochondrial translation initiation (adjusted  $p = 4.43E-09$ ).

Next, GSEA was used to link genes differentially expressed in patients with PBC, PSC, CD, and UC compared with healthy controls and GO terms. Altogether, genes differentially expressed between at least one disease and the control group were attributed to 78 BP, 26 MF, and 23 CC terms (Supplementary Table 4). Of these, all (53 BP, 21 MF, and 15) were in ChLDs, while 35 BP, 21 MF, and 13 CFC terms were identified in IBDs.



**Figure 1.** Venn diagrams illustrating the number of differentially expressed transcripts (adjusted  $p$ -value  $< 0.05$ ) in blood samples from patients with PBC, PSC, and IBD compared with those from healthy controls. PBC; Primary biliary cholangitis, PSC; primary sclerosing cholangitis, IBD; inflammatory bowel disease.

Terms common to ChLDs included 10 BP, 9 MF, and 13 CC terms, and those shared by IBDs comprised 7 BP, 9 MF, and 3 CC terms. Of these, one BP, six MF, and one CC term were common to all four diseases studied, while four BP, two MF, and four CC terms were common to three diseases (Table 3). The majority of terms, along with their child and synonymous terms, which were enriched for differentially expressed genes in one or two of the diseases studied, were related to the endomembrane system, regulation of membrane dynamics by GTPase-mediated processes, and secretion of proinflammatory molecules.

When genes downregulated in blood samples from patients with PBC, PSC, and IBD compared with healthy controls were annotated according to the Reactome database, the following pathways were identified in all three comparisons: R-HSA-163200, Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins; R-HSA-611105, Respiratory electron transport; R-HSA-5389840, Mitochondrial translation elongation; R-HSA-5368286, Mitochondrial translation initiation; R-HSA-5419276, Mitochondrial translation termination; R-HSA-5368287, Mitochondrial translation; R-HSA-1428517, The citric acid (TCA) cycle and respiratory electron transport; and R-HSA-1852241, Organelle biogenesis and maintenance.

**Expression of genes selected for potential use in diagnostic screening.** Among the several hundred probe sets differentially expressed between disease and control groups, the majority exhibited relatively low fold-change (FC) differences in expression level, with no FC values exceeding 1.5 (Fig. 1). To determine whether genes differentially expressed in peripheral blood cells could be used for diagnostic screening, we selected 13 (*EMR1*, *IFI27*, *PLCB2*, *RARA*, *SORL*, *STAT1*, *ABCG1*, *C15orf39*, *LYN*, *PLEKHG3*, *ATG2*, *MME*, *DEFA1*), 15 (*MME*, *FOXO3*, *DBI*, *IFI27*, *HSPE1*, *BOLA2*, *ABCG1*, *PLCB2*, *DYSF*, *CLC*, *PRSS33*, *RAP1*, *GAP*, *RNF182*, *RPS28*), and 7 (*OPLAH*, *ALPL*, *SLC26A8*, *PFKFB3*, *MMP25*, *TLR5*, *DYSF*) genes with expression levels significantly altered in patients with PBC, PSC, and IBDs, respectively, compared with healthy controls. Selected genes were those with differences with the highest level of significance and relatively high FC values and were used for analysis in a confirmation study to determine expression levels by qRT-PCR, using the same RNA samples as those used for microarray profiling. Of the 13, 15, and 7 selected genes, the levels of 7, 7, and 3, respectively, were confirmed to differ significantly (adjusted  $p < 0.05$ ) in samples from patients with PSC, PBC, and IBDs relative to those from healthy control individuals (Table 4).

Next, we assessed the diagnostic potential of all selected genes using an independently recruited cohort of patients and controls. Replication cohorts included 71 patients with PSC, 292 with PBC, and 727 with IBD, along with 206 (PSC, 37; PBC, 138; IBD, 196) controls. The IBD group consisted of 393 patients with CD (253 children and 140 adults) and 334 with UC (199 children and 135 adults). Pair-wise comparisons of qRT-PCR results revealed statistically significant differences (adjusted  $p < 0.05$ ) in expression of five, seven, and six genes between the control group and patients with PSC, PBC, and IBDs, respectively (Table 5).

Next, the diagnostic potential of the mRNAs identified as differentially expressed was assessed using receiver operating characteristic (ROC) curves and area under the curve (AUC) analyses. The AUC-ROC values in PSC, PBC, and IBDs were in the ranges 0.709–0.776, 0.587–0.771, and 0.568–0.650, respectively (Table 5). These values indicate that the tested markers have insufficient discriminatory properties to be applicable for clinical practice. Similar analyses were performed for the CU and UC patient subgroups. AUC-ROC values were in the ranges 0.601–0.682 and 0.575–0.682, respectively, despite highly statistically significant differences in mRNA levels between controls and both the CD and UC subgroups. Furthermore, the highest statistically significant differences for the selected transcripts were obtained for comparisons between active IBDs and controls (range,  $1.36E-07$  to  $1.25E-13$ ). Nevertheless, the corresponding AUC-ROC values were only slightly higher (range,

| ID                        | Description  | qvalue   |
|---------------------------|--|----------|
| <b>Biological process</b> |  |          |
| GO:0042773                | ATP synthesis coupled electron transport   | 1.66E-10 |
| GO:0022904                | respiratory electron transport chain   | 1.66E-10 |
| GO:0042775                | mitochondrial ATP synthesis coupled electron transport                           | 1.66E-10 |
| GO:0022900                | electron transport chain   | 1.66E-10 |
| GO:1902600                | hydrogen ion transmembrane transport   | 3.4E-08  |
| GO:0070125                | mitochondrial translational elongation   | 4.55E-07 |
| GO:0010257                | NADH dehydrogenase complex assembly  | 5.83E-07 |
| GO:0032981                | mitochondrial respiratory chain complex I assembly                               | 5.83E-07 |
| GO:0097031                | mitochondrial respiratory chain complex I biogenesis                             | 5.83E-07 |
| GO:0043087                | regulation of GTPase activity  | 6.04E-05 |
| GO:0043624                | cellular protein complex disassembly   | 0.000260 |
| GO:0000956                | nuclear-transcribed mRNA catabolic process                                       | 0.000294 |
| GO:0022613                | ribonucleoprotein complex biogenesis   | 0.00061  |
| GO:0051056                | regulation of small GTPase mediated signal transduction                          | 0.00168  |
| GO:0045047                | protein targeting to ER  | 0.0057   |
| GO:0019080                | viral gene expression  | 0.0099   |
| GO:0072599                | establishment of protein localization to endoplasmic reticulum                   | 0.0107   |
| GO:0019083                | viral transcription  | 0.0116   |
| GO:0010608                | posttranscriptional regulation of gene expression                                | 0.0265   |
| GO:0002431                | Fc receptor mediated stimulatory signaling pathway                               | 0.033    |
| GO:0006906                | vesicle fusion   | 0.037    |
| GO:0043299                | leukocyte degranulation  | 0.043    |
| GO:0044033                | multi-organism metabolic process   | 0.049    |
| <b>Molecular function</b> |  |          |
| GO:0015078                | hydrogen ion transmembrane transporter activity                                  | 1.72E-09 |
| GO:0003735                | structural constituent of ribosome   | 5.8E-09  |
| GO:0003954                | NADH dehydrogenase activity  | 5.8E-09  |
| GO:0008137                | NADH dehydrogenase (ubiquinone) activity   | 5.8E-09  |
| GO:0050136                | NADH dehydrogenase (quinone) activity  | 5.8E-09  |
| GO:0005096                | GTPase activator activity  | 0.00209  |
| GO:0009055                | electron carrier activity  | 0.0033   |
| GO:0005085                | guanyl-nucleotide exchange factor activity                                       | 0.0150   |
| GO:0044769                | ATPase activity, coupled to transmembrane movement of ions, rotational mechanism | 0.0150   |
| GO:0017137                | Rab GTPase binding   | 0.0150   |
| GO:0015399                | primary active transmembrane transporter activity                                | 0.032    |
| GO:0015405                | P-P-bond-hydrolysis-driven transmembrane transporter activity                    | 0.032    |
| <b>Cellular component</b> |  |          |
| GO:0098800                | inner mitochondrial membrane protein complex                                     | 6.27E-13 |
| GO:0070469                | respiratory chain  | 4.12E-11 |
| GO:0070069                | cytochrome complex   | 0.00071  |
| GO:0099568                | cytoplasmic region   | 0.00201  |
| GO:0030667                | secretory granule membrane   | 0.0117   |
| GO:0061695                | transferase complex, transferring phosphorus-containing groups                   | 0.034    |
| GO:1904115                | axon cytoplasm   | 0.040    |
| GO:0048475                | coated membrane  | 0.048    |

**Table 2.** GO terms over-represented among 1946 probe sets that significantly differentiated disease from control samples in all three comparison groups (i.e., PBC, PSC, and IBD, compared with controls).

0.670–0.739); therefore, our data do not confirm that assessment of levels of these transcripts has discriminatory power to distinguish between samples from patients with disease and healthy controls, even for patients with active intestinal inflammation.

## Discussion

Crosstalk between the gut and the liver may contribute to common mechanisms underlying liver diseases and gastrointestinal and immune disorders. The gut and liver communicate via the biliary tract, portal vein, and systemic circulation<sup>29</sup>; the liver releases BAs and numerous bioactive mediators, while various metabolites produced in the intestine, by both organisms themselves and their gut microbiota, translocate to the liver through the portal vein.

| Biological process |  |     |     |    |    |
|--------------------|--|-----|-----|----|----|
| ID                 | Description  | PBC | PSC | CD | UC |
| GO:0060627         | regulation of vesicle-mediated transport                       | 1   | 1   | 1  | 1  |
| GO:0006906         | vesicle fusion   | 1   | 1   | 1  | 0  |
| GO:0016050         | vesicle organization   | 1   | 1   | 0  | 1  |
| GO:0031346         | positive regulation of cell projection organization            | 1   | 1   | 1  | 0  |
| GO:0032479         | regulation of type I interferon production                     | 1   | 1   | 1  | 0  |
| Molecular function |  |     |     |    |    |
| ID                 | Description  | PBC | PSC | CD | MF |
| GO:0003682         | chromatin binding  | 1   | 1   | 1  | 1  |
| GO:0004674         | protein serine/threonine kinase activity                       | 1   | 1   | 1  | 1  |
| GO:0005096         | GTPase activator activity                                      | 1   | 1   | 1  | 1  |
| GO:0004386         | helicase activity  | 1   | 1   | 1  | 1  |
| GO:0005085         | guanyl-nucleotide exchange factor activity                     | 1   | 1   | 1  | 1  |
| GO:0017016         | Ras GTPase binding   | 1   | 1   | 1  | 1  |
| GO:0005543         | phospholipid binding   | 1   | 1   | 0  | 1  |
| GO:0003779         | actin binding  | 1   | 1   | 1  | 0  |
| GO:0022804         | active transmembrane transporter activity                      | 1   | 1   | 0  | 1  |
| Cellular component |  |     |     |    |    |
| ID                 | Description  | PBC | PSC | CD | UC |
| GO:0061695         | transferase complex, transferring phosphorus-containing groups | 1   | 1   | 1  | 1  |
| GO:0005924         | cell-substrate adherens junction                               | 1   | 1   | 0  | 1  |
| GO:0005765         | lysosomal membrane   | 1   | 1   | 0  | 1  |
| GO:0030667         | secretory granule membrane                                     | 1   | 1   | 1  | 0  |
| GO:0015629         | actin cytoskeleton   | 1   | 1   | 1  | 0  |

**Table 3.** GO terms significantly associated with changes in gene expression between control and disease samples according to GSEA analysis.

**Functional analysis of WBC gene expression profiles across PBC, PCS, and IBDs.** High-density microarrays allow the measurement of gene expression without prior knowledge of expression profiles. Expression profiles repeatedly measured in whole blood samples from healthy subjects generate repeatable data, from each individual subject, over several months<sup>30</sup>. Specific profiles associated with affected status have been identified in a wide range of diseases, including autoimmune and inflammatory diseases, infectious disorders, psychiatric, cardiovascular, neurological, and neoplastic diseases, and even various environmental factors<sup>31,32</sup>. Among associated environmental factors, blood transcriptome variables could identify associations of socioeconomic status with chronic inflammation<sup>33–35</sup> and exhibited species- and strain-level specificity in discrimination of viral, bacterial, and eukaryotic infectious diseases, including acute and chronic active Epstein–Barr virus infection and response to tuberculosis treatment<sup>36</sup>. Predictive biomarkers in peripheral blood samples can identify patients with intracranial aneurysms<sup>37</sup>, be used to stratify patients according to disease progression before and after the onset of type 1 diabetes<sup>38–41</sup>, and classify systemic lupus erythematosus and rheumatoid arthritis by prediction of their responsiveness to anti-IFN therapy<sup>42,43</sup>. A few studies have also described alterations of WBC gene expression profiles in IBDs<sup>18–21</sup>.

In this study, we evaluated the molecular alterations underlying PBC, PSC, and IBDs, by functional analysis of microarray data sets through annotation according to the GO and Reactome databases. The majority of terms extracted, based on enrichment for genes differentially expressed in pair-wise comparisons between healthy controls and patients with PBC, PSC, and IBD, shared common profiles related to the vesicle endomembrane system and GTPase-mediated processes. A second major group of GO terms attributed to probe sets with expression changes in all three disease types (PBC, PSC, and IBDs) were related to mitochondrial function. Overall, these terms represent immunological and inflammatory pathways related to cellular stress. Similar functional alterations in WBC transcriptomes were also reported in many of the conditions mentioned above.

Dysregulation of innate and adaptive immune processes is associated with both IBDs and autoimmune fibrous cholangiopathies<sup>6,44–52</sup>. The epithelium of the gastrointestinal tract forms a physical barrier against microbes, and Paneth and goblet cells monitor the bacterial community and regulate host–microbe homeostasis through the production of antimicrobial peptides and mucins. Once the intestinal defense system is affected, or the ecological organization of the healthy gut microbiota is disturbed, immune and inflammatory responses are activated, and can lead to the accumulation of ROS, endoplasmic reticulum (ER) stress, and mitochondrial dysfunction<sup>53,54</sup>. Gut dysbiosis may also be related to alterations in BAs; increased concentrations of hydrophobic BAs may lead to mitochondrial and ER stress-related activation of death receptors and production of inflammatory mediators, such as cytokines, chemokines, and adhesion molecules. Overall, such changes can initiate cholangiocyte cytotoxicity; therefore, the BA–intestinal microbiota–cholestasis triangle is postulated to play a vital role in the pathogenesis of PBC and PSC<sup>44</sup>.

| PSC    | p value         | FC   | AUC   | PBC      | p value         | FC   | AUC   | IBDs    | p value         | FC   | AUC   |
|--------|-----------------|------|-------|----------|-----------------|------|-------|---------|-----------------|------|-------|
| MME    | <b>0.00039</b>  | 1.3  | 0.709 | EMR1     | <b>1.67E-06</b> | 1.29 | 0.643 | OPLAH   | <b>1.13E-05</b> | 1.55 | 0.603 |
| FOXO3  | <b>0.000267</b> | 1.5  | 0.715 | IFI27    | <b>1.69E-07</b> | 4.51 | 0.656 | ALPL    | <b>9.34E-09</b> | 1.61 | 0.650 |
| ABCG1  | <b>2.78E-06</b> | 0.63 | 0.776 | PLCB2    | 0.003413        | 1.14 | 0.587 | SLC26A8 | <b>1.01E-10</b> | 1.80 | 0.568 |
| CLC    | <b>0.000136</b> | 1.97 | 0.725 | RARA     | <b>0.00016</b>  | 1.17 | 0.612 | PFKFB3  | <b>0.003463</b> | 1.30 | 0.601 |
| PRSS33 | <b>0.000311</b> | 2.21 | 0.715 | STAT1    | <b>7.56E-20</b> | 1.42 | 0.771 | MMP25   | <b>1.8E-05</b>  | 1.41 | 0.600 |
|        |                 |      |       | ABCG1    | <b>1.43E-08</b> | 1.26 | 0.669 | TLR5    | <b>1.73E-10</b> | 1.45 | 0.648 |
|        |                 |      |       | C15orf39 | <b>1.26E-05</b> | 1.33 | 0.630 | DYSF    | <b>1.47E-07</b> | 1.53 | 0.622 |

**Table 4.** Results of confirmation analysis of selected gene expression differences by qRT-PCR. FC, fold-change; AUC, area under the curve.

| Gene    | Crohn's disease |      |       | Ulcerative colitis |      |       | Active IBDs     |      |       |
|---------|-----------------|------|-------|--------------------|------|-------|-----------------|------|-------|
|         | P value         | FC   | AUC   | P value            | FC   | AUC   | P value         | FC   | AUC   |
| OPLAH   | <b>5.98E-07</b> | 1.53 | 0.628 | <b>0.0037</b>      | 1.57 | 0.576 | <b>1.03E-08</b> | 2.49 | 0.686 |
| ALPL    | <b>5.44E-11</b> | 1.63 | 0.666 | <b>0.0002</b>      | 1.59 | 0.597 | <b>2.2E-09</b>  | 2.25 | 0.693 |
| SLC26A8 | <b>2.41E-13</b> | 1.93 | 0.686 | <b>2E-05</b>       | 1.66 | 0.611 | <b>1.47E-12</b> | 2.68 | 0.728 |
| PFKFB3  | <b>7.2E-05</b>  | 1.30 | 0.601 | 0.2192             | 1.30 | 0.601 | <b>1.21E-09</b> | 1.82 | 0.696 |
| MMP25   | <b>1.38E-07</b> | 1.45 | 0.633 | 0.0162             | 1.36 | 0.633 | <b>1.36E-07</b> | 1.86 | 0.670 |
| TLR5    | <b>6.46E-13</b> | 1.51 | 0.682 | <b>2E-05</b>       | 1.38 | 0.682 | <b>1.25E-13</b> | 2.06 | 0.739 |
| DYSF    | <b>7.63E-11</b> | 1.63 | 0.665 | <b>0.0038</b>      | 1.43 | 0.575 | <b>5.92E-10</b> | 2.06 | 0.699 |

**Table 5.** Results of replication analysis of selected gene expression differences by qRT-PCR using samples from an independent cohort. FC, fold-change; AUC, area under the curve.

The mechanisms underlying autoimmune liver diseases and gastrointestinal disorders are associated with recirculation of the cell membrane. Exosome vesicles packed with bioactive molecules are involved in cytokine secretion and adaptive immune responses<sup>55,56</sup> and act as mediators between neighboring cells and distant organs<sup>57,58</sup>. The intracellular transport and delivery of vesicles to the plasma membrane involves GTP-binding proteins<sup>59</sup> and depends on actin cytoskeleton organization, which dynamically regulates directed endosome traffic and recycling involved in the immune and stress responses<sup>60,61</sup>. Autophagy, an effector mechanism of cellular senescence that blocks the proliferation of cells that harbor genomic injuries, is a lysosome-dependent protective response against various cellular stresses. Autophagy involves degradation and recycling of protein aggregates and damaged organelles and is pivotal for secretion of proteins and production of antimicrobial peptides. The autophagy process regulates a number of cellular functions, including inflammation and adaptive immunity, host defenses, mitochondrial homeostasis, and lipid metabolism, and controls the balance between abnormal immune activation and inflammation<sup>53,54,62–65</sup>.

Finally, the majority of GO nodes extracted from blood transcriptomes were common to phenotypically dissimilar disorders, including ChLDs and IBDs, and were consistent with previous studies uncovering alterations of WBC gene expression in IBDs<sup>18–21</sup>.

**The diagnostic utility of screening for expression of selected WBC genes in PBC, PCS, and IBDs.** Gene expression microarray technology can be used to identify genes that are differentially expressed between predefined groups of samples (class comparison), genes whose expression differs across predefined classes of genes (class prediction), and genes that allow classification of molecular subgroups among individuals with seemingly homogenous phenotypes (class discovery). The final results of expression profiling consist of lists of measurements directly linked to genes, some of which may be used as diagnostic, prognostic, or predictive biomarkers. Biomarkers are typically identified by high-throughput methods and subsequently validated by standard molecular methods. In this study, the selection of potential biomarkers was conducted using microarray profiling of gene expression and, since microarray data typically exhibit a low degree of reproducibility<sup>66</sup>, the selected measurements were directly verified by confirmation analysis and indirectly confirmed by qRT-PCR replication studies.

Our microarray-based studies identified thousands of probe sets that differed between disease and control samples; however, the majority of these exhibited low FC values. As higher FC values are positively correlated with the probability that a biomarker can meet the expectations required for clinical utility, we selected genes exhibiting the most statistically significant and largest FC differences between patient and control samples. Although the FC values of the majority of selected genes did not exceed two, both confirmation and replication studies demonstrated that some of them exhibited significant differences in expression between the disease and control groups, with the highest level of significance in patients with active IBDs (p-value range, 1.36E-07 to 1.25E-13). Additionally, we found that 86 differentially expressed genes from our study were common with a set of 133 genes that were designated by Peters *et al.*<sup>67</sup> as the key driver genes of IBD (Supplementary Fig. S3). Of these 37 were shared among the diseases and 15, 7, and 3 were unique for IBD, PBC and PSC, respectively. This extensive overlap again indicates a functional link between IBD susceptibility genes expression contributing to a discrete systemic inflammation that can be portrayed in blood transcriptome.

Numerous previous studies reported the clinical utility of blood RNA expression profiles; however, many did not perform further validation experiments to demonstrate the utility of their assays for clinical diagnosis. Although medical classification should ideally be binary, i.e., dividing a population by the presence or absence of disease, the majority of molecular biomarkers generate results that overlap between health and disease states. Consequently, most so-called biomarkers can discriminate between groups of patients and controls, rather than being able to consistently and completely distinguish individuals with, from those without, a disease of interest. AUC-ROC values are an appropriate means of assessing the relationship between the sensitivity and specificity of a biomarker across all potential cut-off values. AUC-ROC values  $>0.8$  are assumed to represent moderate (good) discriminatory power, with those  $>0.9$  considered to indicate high (excellent) power to distinguish between analyzed groups. Unexpectedly, according to the AUC-ROCs calculated based on qRT-PCR analysis of expression levels in this study, no single RNA reached diagnostic potential. Our results are consistent with the AUC-ROC values calculated for changes in blood transcriptional levels determined by monitoring UC patients over time in a previous study, which did not exceed 0.8<sup>20</sup>; however, they differ from the results of a recently published study reporting a panel of six genes that could distinguish CD and UC with AUC-ROCs ranging from 0.89 to 0.99<sup>19</sup>. In the latter study, the predictive performance was based on PCR data from only 20 samples<sup>19</sup>. Indeed, blood expression profiles have previously been examined in rather small populations, and analyses of differentially expressed genes have generally produced results with overlap between healthy and diseased samples<sup>18–21</sup>. Our microarray screening, followed by confirmatory qRT-PCR studies, was conducted using 370 RNA samples, and several hundred additional samples were included in the replication analysis. Therefore, the results of our investigation can be considered reliable, since the approach applied was appropriate for a search for new biomarkers and employed a relatively large patient population.

To summarize, although we are witnessing the era of molecular diagnostics, of the numerous potential biomarkers identified by high-throughput methods in chronic autoimmune diseases, none has proven ideal to date<sup>68,69</sup>. This study indicates that microarray-based profiling of blood gene expression levels can support research into the molecular mechanisms underlying disease, while being less useful for the selection of diagnostic biomarkers for use in clinical practice.

### Data Availability

The results of microarray measurements have been deposited in Gene Expression Omnibus database, entry GSE119600.

### References

- Webb, G. J. & Hirschfield, G. M. Using GWAS to identify genetic predisposition in hepatic autoimmunity. *J. Autoimmun.* **66**, 25–39 (2016).
- McGovern, D. P. B., Kugathasan, S. & Cho, J. H. Genetics of Inflammatory Bowel Diseases. *Gastroenterology* **149**, 1163–1176.e2 (2015).
- de Souza, H. S. P. & Fiocchi, C. Immunopathogenesis of IBD: current state of the art. *Nat. Rev. Gastroenterol. Hepatol.* **13**, 13–27 (2016).
- Gulamhusein, A. F., Juran, B. D. & Lazaridis, K. N. Genome-Wide Association Studies in Primary Biliary Cirrhosis. *Semin. Liver Dis.* **35**, 392–401 (2015).
- Hirschfield, G. M. & Gershwin, M. E. The immunobiology and pathophysiology of primary biliary cirrhosis. *Annu. Rev. Pathol.* **8**, 303–330 (2013).
- Nakanuma, Y., Sasaki, M. & Harada, K. Autophagy and senescence in fibrosing cholangiopathies. *J. Hepatol.* **62**, 934–945 (2015).
- Karlsen, T. H. *et al.* Genome-wide association analysis in primary sclerosing cholangitis. *Gastroenterology* **138**, 1102–1111 (2010).
- Melum, E. *et al.* Genome-wide association analysis in primary sclerosing cholangitis identifies two non-HLA susceptibility loci. *Nat. Genet.* **43**, 17–19 (2011).
- Folseraas, T. *et al.* Extended analysis of a genome-wide association study in primary sclerosing cholangitis detects multiple novel risk loci. *J. Hepatol.* **57**, 366–375 (2012).
- Ellinghaus, D. *et al.* Genome-wide association analysis in primary sclerosing cholangitis and ulcerative colitis identifies risk loci at GPR35 and TCF4. *Hepatol. Baltim. Md* **58**, 1074–1083 (2013).
- Boonstra, K. *et al.* Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology* **58**, 2045–2055 (2013).
- Anderson, C. A. *et al.* Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat. Genet.* **43**, 246–252 (2011).
- Janse, M. *et al.* Three ulcerative colitis susceptibility loci are associated with primary sclerosing cholangitis and indicate a role for IL2, REL, and CARD9. *Hepatol. Baltim. Md* **53**, 1977–1985 (2011).
- Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119–124 (2012).
- Ruel, J., Ruane, D., Mehandru, S., Gower-Rousseau, C. & Colombel, J.-F. IBD across the age spectrum: is it the same disease? *Nat. Rev. Gastroenterol. Hepatol.* **11**, 88–98 (2014).
- Liew, C.-C., Ma, J., Tang, H.-C., Zheng, R. & Dempsey, A. A. The peripheral blood transcriptome dynamically reflects system wide biology: a potential diagnostic tool. *J. Lab. Clin. Med.* **147**, 126–132 (2006).
- European Association for the Study of the Liver. EASL Clinical Practice Guidelines: management of cholestatic liver diseases. *J. Hepatol.* **51**, 237–267 (2009).
- Barnes, E. L., Liew, C.-C., Chao, S. & Burakoff, R. Use of blood based biomarkers in the evaluation of Crohn's disease and ulcerative colitis. *World J. Gastrointest. Endosc.* **7**, 1233–1237 (2015).
- Burakoff, R. *et al.* Blood-based biomarkers used to predict disease activity in Crohn's disease and ulcerative colitis. *Inflamm. Bowel Dis.* **21**, 1132–1140 (2015).
- Planell, N. *et al.* Usefulness of Transcriptional Blood Biomarkers as a Non-invasive Surrogate Marker of Mucosal Healing and Endoscopic Response in Ulcerative Colitis. *J. Crohns Colitis* **11**, 1335–1346 (2017).
- Burakoff, R. *et al.* Differential regulation of peripheral leukocyte genes in patients with active Crohn's disease and Crohn's disease in remission. *J. Clin. Gastroenterol.* **44**, 120–126 (2010).
- Best, W. R., Beckett, J. M., Singleton, J. W. & Kern, F. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. *Gastroenterology* **70**, 439–444 (1976).
- Seo, M. *et al.* An index of disease activity in patients with ulcerative colitis. *Am. J. Gastroenterol.* **87**, 971–976 (1992).

24. Hyams, J. S. *et al.* Development and validation of a pediatric Crohn's disease activity index. *J. Pediatr. Gastroenterol. Nutr.* **12**, 439–447 (1991).
25. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).
26. Mikula, M. *et al.* Integrating proteomic and transcriptomic high-throughput surveys for search of new biomarkers of colon tumors. *Funct. Integr. Genomics* **11**, 215–224 (2011).
27. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-ΔΔC<sub>T</sub></sup> Method. *Methods San Diego Calif* **25**, 402–408 (2001).
28. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic J. Integr. Biol.* **16**, 284–287 (2012).
29. Tripathi, A. *et al.* The gut–liver axis and the intersection with the microbiome. *Nat. Rev. Gastroenterol. Hepatol.* **1**, <https://doi.org/10.1038/s41575-018-0011-z> (2018).
30. De Boever, P. *et al.* Characterization of the peripheral blood transcriptome in a repeated measures design using a panel of healthy individuals. *Genomics* **103**, 31–39 (2014).
31. Cabrera, S. M., Chen, Y.-G., Hagopian, W. A. & Hessner, M. J. Blood-based signatures in type 1 diabetes. *Diabetologia* **59**, 414–425 (2016).
32. Mesko, B., Poliska, S. & Nagy, L. Gene expression profiles in peripheral blood for the diagnosis of autoimmune diseases. *Trends Mol. Med.* **17**, 223–233 (2011).
33. Gaye, A., Gibbons, G. H., Barry, C., Quarells, R. & Davis, S. K. Influence of socioeconomic status on the whole blood transcriptome in African Americans. *PLoS One* **12**, e0187290 (2017).
34. Chen, E., Fisher, E. B., Bacharier, L. B. & Strunk, R. C. Socioeconomic status, stress, and immune markers in adolescents with asthma. *Psychosom. Med.* **65**, 984–992 (2003).
35. Piers, L. H. *et al.* Relation of aortic valve and coronary artery calcium in patients with chronic kidney disease to the stage and etiology of the renal disease. *Am. J. Cardiol.* **103**, 1473–1477 (2009).
36. Blankley, S. *et al.* The application of transcriptional blood signatures to enhance our understanding of the host response to infection: the example of tuberculosis. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **369**, 20130427 (2014).
37. Tutino, V. M. *et al.* Circulating neutrophil transcriptome may reveal intracranial aneurysm signature. *PLoS One* **13**, e0191407 (2018).
38. Jin, Y. *et al.* Risk of type 1 diabetes progression in islet autoantibody-positive children can be further stratified using expression patterns of multiple genes implicated in peripheral blood lymphocyte activation and function. *Diabetes* **63**, 2506–2515 (2014).
39. Irvine, K. M. *et al.* Peripheral blood monocyte gene expression profile clinically stratifies patients with recent-onset type 1 diabetes. *Diabetes* **61**, 1281–1290 (2012).
40. Reynier, F. *et al.* Specific gene expression signature associated with development of autoimmune type-1 diabetes using whole-blood microarray analysis. *Genes Immun.* **11**, 269–278 (2010).
41. Kallionpää, H. *et al.* Innate immune activity is detected prior to seroconversion in children with HLA-conferred type 1 diabetes susceptibility. *Diabetes* **63**, 2402–2414 (2014).
42. Pascual, V., Chaussabel, D. & Banchereau, J. A genomic approach to human autoimmune diseases. *Annu. Rev. Immunol.* **28**, 535–571 (2010).
43. Kirou, K. A. & Gkrouzman, E. Anti-interferon alpha treatment in SLE. *Clin. Immunol. Orlando Fla* **148**, 303–312 (2013).
44. Li, Y., Tang, R., Leung, P. S. C., Gershwin, M. E. & Ma, X. Bile acids and intestinal microbiota in autoimmune cholestatic liver diseases. *Autoimmun. Rev.* **16**, 885–896 (2017).
45. Allen, K., Jaeschke, H. & Copple, B. L. Bile acids induce inflammatory genes in hepatocytes: a novel mechanism of inflammation during obstructive cholestasis. *Am. J. Pathol.* **178**, 175–186 (2011).
46. Tilg, H., Cani, P. D. & Mayer, E. A. Gut microbiome and liver diseases. *Gut* **65**, 2035–2044 (2016).
47. Shamriz, O. *et al.* Microbiota at the crossroads of autoimmunity. *Autoimmun. Rev.* **15**, 859–869 (2016).
48. Rossen, N. G. *et al.* The mucosa-associated microbiota of PSC patients is characterized by low diversity and low abundance of uncultured Clostridiales II. *J. Crohns Colitis* **9**, 342–348 (2015).
49. Lv, L.-X. *et al.* Alterations and correlations of the gut microbiome, metabolism and immunity in patients with primary biliary cirrhosis. *Environ. Microbiol.* **18**, 2272–2286 (2016).
50. Tang, R. *et al.* Gut microbial profile is altered in primary biliary cholangitis and partially restored after UDCA therapy. *Gut* **67**, 534–541 (2018).
51. Kostic, A. D., Xavier, R. J. & Gevers, D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* **146**, 1489–1499 (2014).
52. Winter, S. E., Lopez, C. A. & Bäumlér, A. J. The dynamics of gut-associated microbial communities during inflammation. *EMBO Rep.* **14**, 319–327 (2013).
53. Wang, S.-L. *et al.* Impact of Paneth Cell Autophagy on Inflammatory Bowel Disease. *Front. Immunol.* **9**, 693 (2018).
54. Coleman, O. I. & Haller, D. Bacterial Signaling at the Intestinal Epithelial Interface in Inflammation and Cancer. *Front. Immunol.* **8**, 1927 (2017).
55. Soares, H. *et al.* Regulated vesicle fusion generates signaling nanoterritories that control T cell activation at the immunological synapse. *J. Exp. Med.* **210**, 2415–2433 (2013).
56. Bustos-Morán, E., Blas-Rus, N., Martín-Cófreces, N. B. & Sánchez-Madrid, F. Microtubule-associated protein-4 controls nanovesicle dynamics and T cell activation. *J. Cell Sci.* **130**, 1217–1223 (2017).
57. Nawaz, M. Extracellular vesicle-mediated transport of non-coding RNAs between stem cells and cancer cells: implications in tumor progression and therapeutic resistance. *Stem Cell Investig.* **4**, 83 (2017).
58. Fung, K. Y. Y., Fairn, G. D. & Lee, W. L. Transcellular vesicular transport in epithelial and endothelial cells: Challenges and opportunities. *Traffic Cph. Den.* **19**, 5–18 (2018).
59. Watson, E. L. GTP-binding proteins and regulated exocytosis. *Crit. Rev. Oral Biol. Med. Off. Publ. Am. Assoc. Oral Biol.* **10**, 284–306 (1999).
60. Hafner, A. E. & Rieger, H. Spatial Cytoskeleton Organization Supports Targeted Intracellular Transport. *Biophys. J.* **114**, 1420–1432 (2018).
61. Vertii, A., Hehnlly, H. & Doxsey, S. The Centrosome, a Multitalented Renaissance Organelle. *Cold Spring Harb. Perspect. Biol.* **8** (2016).
62. Pavel, M. & Rubinsztein, D. C. Mammalian autophagy and the plasma membrane. *FEBS J.* **284**, 672–679 (2017).
63. Lu, Y., Li, X., Liu, S., Zhang, Y. & Zhang, D. Toll-like Receptors and Inflammatory Bowel Disease. *Front. Immunol.* **9**, 72 (2018).
64. Kabat, A. M., Pott, J. & Maloy, K. J. The Mucosal Immune System and Its Regulation by Autophagy. *Front. Immunol.* **7**, 240 (2016).
65. Lee, H.-Y. *et al.* Autophagy deficiency in myeloid cells increases susceptibility to obesity-induced diabetes and experimental colitis. *Autophagy* **12**, 1390–1403 (2016).
66. Ostrowski, J. & Wyrwicz, L. S. Integrating genomics, proteomics and bioinformatics in translational studies of molecular medicine. *Expert Rev. Mol. Diagn.* **9**, 623–630 (2009).
67. Peters, L. A. *et al.* A functional genomics predictive network model identifies regulators of inflammatory bowel disease. *Nat. Genet.* **49**, 1437–1449 (2017).

68. Viennois, E., Zhao, Y. & Merlin, D. Biomarkers of IBD: from classical laboratory tools to personalized medicine. *Inflamm. Bowel Dis.* **21**, 2467–2474 (2015).
69. Olaizola, P. *et al.* MicroRNAs and extracellular vesicles in cholangiopathies. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* **1864**, 1293–1307 (2018).

## Acknowledgements

This work was supported by: the National Science Centre [2011/01/B/NZ5/05291 and 2011/02/A/NZ5/00339].

## Author Contributions

Conception and design of the study: J.O. and A.H. Patients recruitment and clinical data compilation: J.O., A.H., I.L. and A.R. The Polish IBD study Group, The Polish PBC study Group; RNA isolation, RNA quality control, RT-qPCR: A.P., F.A., N.Z.L, J.K. A.K., M.P., A.B., P.M., A.D. and M.D. dataset analyses and interpretation: J.O., K.G., M.M. and M.K. Drafting of the manuscript: J.O., M.M. and K.G.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-43699-1>.

**Competing Interests:** Prof. Ostrowski and prof. Habor report grants from The National Science Centre, during the conduct of the study. The funders had no role in the study design, data collection, analysis and interpretation, decision to publish, or preparation of the manuscript. Other authors report no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019

## Consortia

### The Polish PBC study Group

Bozena Walewska-Zielecka<sup>4</sup>, Marek Krawczyk<sup>5</sup>, Halina Cichoz-Lach<sup>6</sup>, Piotr Milkiewicz<sup>7,8</sup>, Agnieszka Kowalik<sup>6</sup>, Krzysztof Mucha<sup>9,10</sup>, Joanna Raczynska<sup>9</sup>, Joanna Musialik<sup>11</sup>, Grzegorz Boryczka<sup>11</sup>, Michal Wasilewicz<sup>7</sup>, Irena Ciecko-Michalska<sup>12</sup>, Malgorzata Ferenc<sup>13</sup>, Maria Janiak<sup>14</sup>, Alina Kanikowska<sup>15</sup>, Rafal Stankiewicz<sup>5</sup>, Marek Hartleb<sup>11</sup>, Tomasz Mach<sup>12</sup>, Marian Grzymislowski<sup>15</sup>, Joanna Raszeja-Wyszomirska<sup>7</sup>, Ewa Wunsch<sup>8</sup> & Tomasz Bobinski<sup>16</sup>

### The Polish IBD study Group

Jaroslaw Kierkus<sup>17</sup>, Piotr Socha<sup>17</sup>, Michal Lodyga<sup>18</sup>, Maria Klopocka<sup>19</sup>, Barbara Iwanczak<sup>20</sup>, Katarzyna Bak-Drabik<sup>21</sup>, Jaroslaw Walkowiak<sup>22</sup>, Piotr Radwan<sup>23</sup>, Urszula Grzybowska-Chlebowczyk<sup>24</sup>, Bartosz Korczowski<sup>25</sup> & Teresa Starzynska<sup>26</sup>

<sup>4</sup>Department of Public Health, Faculty of Health Sciences, Medical University of Warsaw, Warsaw, Poland.

<sup>5</sup>Department of General, Transplant and Liver Surgery, Medical University of Warsaw, Warsaw, Poland. <sup>6</sup>Department of Gastroenterology, Medical University, Lublin, Poland. <sup>7</sup>Department of General, Liver and Internal Medicine Unit,

Transplant and Liver Surgery, Medical University of Warsaw, Warsaw, Poland. <sup>8</sup>Department of Clinical and Molecular

Biochemistry, Pomeranian Medical University, Szczecin, Poland. <sup>9</sup>Department of Immunology, Transplantology and Internal Medicine, Medical University of Warsaw, Warsaw, Poland. <sup>10</sup>Institute of Biochemistry and Biophysics, Polish

Academy of Sciences, Warsaw, Poland. <sup>11</sup>Department of Gastroenterology and Hepatology, Medical University of Silesia, Katowice, Poland. <sup>12</sup>Department of Gastroenterology and Infectious Diseases, Collegium Medicum

Jagiellonian University, Krakow, Poland. <sup>13</sup>Department of Gastroenterology, Provincial Hospital, Olsztyn, Poland.

<sup>14</sup>Department of Gastroenterology and Hepatology, Medical University of Gdansk, Gdansk, Poland. <sup>15</sup>Department of

Internal and Metabolic Diseases and Dietetics, Poznan University of Medical Sciences, Poznan, Poland. <sup>16</sup>Department of Gastroenterology, Provincial Hospital, Ostroleka, Poland. <sup>17</sup>Department of Gastroenterology, Hepatology and

Feeding Disorders, Children's Memorial Health Institute, Warsaw, 04-730, Poland. <sup>18</sup>Department of Internal Medicine and Gastroenterology with IBD Subdivision, Central Clinical Hospital of the Ministry of the Interior, Warsaw, 02-507,

Poland. <sup>19</sup>Vascular Diseases and Internal Medicine, Nicolaus Copernicus University in Torun, Collegium Medicum, Bydgoszcz, 85-067, Poland. <sup>20</sup>Department of Pediatrics, Gastroenterology and Nutrition, Wroclaw Medical University,

Wroclaw, 50-367, Poland. <sup>21</sup>Department of Pediatrics, School of Medicine with the Division of Dentistry in Zabrze, Medical University of Silesia, Katowice, 40-752, Poland. <sup>22</sup>Department of Pediatric Gastroenterology & Metabolic

Diseases, Poznan University of Medical Sciences, Poznan, 61-701, Poland. <sup>23</sup>Department of Gastroenterology, Medical University of Lublin, Lublin, 20-059, Poland. <sup>24</sup>Department of Pediatrics, School of Medicine in Katowice,

Medical University of Silesia, Katowice, 40-752, Poland. <sup>25</sup>Medical College, University of Rzeszow, Rzeszow, 35-959, Poland. <sup>26</sup>Department of Gastroenterology, Pomeranian Medical University, Szczecin, 70-204, Poland.