Cross-species high-resolution transcriptome profiling suggests biomarkers and therapeutic targets for ulcerative colitis

Reza Yarani[1]*[†], Oana Palasca[2,3,4†], Nadezhda T. Doncheva[2,3,4], Christian Anthon[3,4], Bartosz Pilecki[5], Cecilie A. S. Svane[1], Aashiq H. Mirza[3,6], Thomas Litman[7], Uffe Holmskov[5], Claus H. Bang-Berthelsen[8,9], Mogens Vilien[10], Lars J. Jensen[2,3], Jan Gorodkin[3,4] and Flemming Pociot[1,3,11,12]*

[1]Translational Type 1 Diabetes Research, Department of Clinical Research, Steno Diabetes Center Copenhagen, Gentofte, Denmark, [2]Novo Nordisk Foundation Center for Protein Research, University of Copenhagen, Copenhagen, Denmark, [3]Center for non-coding RNA in Technology and Health, University of Copenhagen, Copenhagen, Denmark, [4]Department of Veterinary and Animal Sciences, University of Copenhagen, Copenhagen, Denmark, [5]Department of Cancer and Inflammation Research, Institute of Molecular Medicine, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark, [6]Department of Pharmacology, Weill Cornell Medicine, Cornell University, New York, NY, United States, [7]Department of Immunology and Microbiology, University of Copenhagen, Copenhagen, Denmark, [8]Research Group for Microbial Biotechnology and Biorefining, National Food Institute, Technical University of Denmark, Kgs. Lyngby, Denmark, [9]Department of Gastroenterology, North Zealand Hillerød Hospital, Hillerød, Denmark, [10]Department of Surgery, North Zealand Hospital, Hillerød, Denmark, [11]Copenhagen Diabetes Research Center, Department of Pediatrics, Herlev University Hospital, Herlev, Denmark, [12]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Background: Ulcerative colitis (UC) is a disorder with unknown etiology, and animal models play an essential role in studying its molecular pathophysiology. Here, we aim to identify common conserved pathological UC-related gene expression signatures between humans and mice that can be used as treatment targets and/or biomarker candidates.

Methods: To identify differentially regulated protein-coding genes and non-coding RNAs, we sequenced total RNA from the colon and blood of the most widely used dextran sodium sulfate Ulcerative colitis mouse. By combining this with public human Ulcerative colitis data, we investigated conserved gene expression signatures and pathways/biological processes through which these genes may contribute to disease development/progression.

Results: Cross-species integration of human and mouse Ulcerative colitis data resulted in the identification of 1442 genes that were significantly differentially regulated in the same direction in the colon and 157 in blood. Of these, 51 genes showed consistent differential regulation in the colon and blood. Less known genes with importance in disease pathogenesis, including *SPI1, FPR2, TYROBP, CKAP4, MCEMP1, ADGRG3, SLC11A1,* and *SELPLG,* were identified through network centrality ranking and validated in independent human and mouse cohorts.

Conclusion: The identified Ulcerative colitis conserved transcriptional signatures aid in the disease phenotyping and future treatment decisions, drug discovery, and clinical trial design.

Introduction

Ulcerative colitis (UC) is an inflammatory bowel disorder mainly affecting the large intestine. Most often, UC initiates from the rectum, affects the mucosal lining, and is limited to the innermost layers of the large intestine. Multiple pathogenic factors, including numerous susceptibility gene variants, environmental factors, changes in the gut microbiota, and a dysregulated immune response, are all associated with UC. Despite this recognition and the identification of apparently relevant factors, a complete understanding of UC pathogenesis still needs to be reached, and thus, treatment may not be optimal. An important reason for this unsatisfactory situation is the currently limited comprehension of the genuinely relevant components of UC immuno-pathogenesis. Thus, given the complex nature of UC, the study of animal models is crucial. A big challenge in using animal models for UC is whether the model's genes and molecular pathways are analogous to humans. Although many coding and non-coding genes between humans and mice are highly similar (Hong et al., 2017), they might be regulated differently, and the pathophysiology may differ. The expression of both differentially regulated coding and non-coding genes contributes to disease pathology (Casamassimi et al., 2017; Lekka and Hall, 2018). Differentially regulated long non-coding RNAs (lncRNAs) and microRNAs (miRNAs) are linked to the change in the expression of many protein-coding genes (PCGs) and, subsequently, the development of disease. Many studies are made without taking the non-coding RNAs, i.e., disease-related pathways and differentially expressed genes, into account, and thereby, fundamental and essential players in the disease regulatory networks may be overlooked (Hasler et al., 2017).

Preclinical murine models have been used for *in vivo* assessments of UC. One of the best-established UC models is the dextran sulfate sodium (DSS) model (Chassaing et al., 2014; Eichele and Kharbanda, 2017). DSS is a negatively charged polysaccharide with colitogenic properties that induce colitis when given orally (Chassaing et al., 2014). Although the mechanisms by which DSS causes intestinal inflammation are not fully understood (Okayasu et al., 1990), it seems that the exposed colonic monolayer of epithelial cells in the large intestine, specifically the distal colon, are being progressively eroded, and the lamina propria barrier integrity is compromised, in which an enormous number of microorganisms live. Therefore, inflammation is triggered following bacteria (and their products) translocation across the damaged intestinal wall. However, to our knowledge, a comprehensive investigation of the DSS model's coding and non-coding colonic and blood transcriptome and its comparison with human UC has not been performed as of today.

We, therefore, set out to characterize the transcriptomic landscape of the DSS-UC mouse model by performing transcriptomic profiling of PCGs, lncRNAs, and miRNAs from colon and whole blood using deep RNA sequencing. We identified differentially expressed lncRNAs and underscored their functional importance by annotating their expressed neighboring PCGs. Then we obtained the targets for the differentially regulated miRNAs and identified the biological processes that these targets are affecting. Further, we assessed the similarity between the experimental model and human UC. To this aim, several high-quality publicly available human UC colonic and blood transcriptomic studies, including our previous study, were selected, data obtained, processed, and cross-species UC transcriptomic landscape was compared. Our results highlight widespread differential regulation of coding and non-coding genes in the mouse and human UC models with marked similarities in both tissues, specifically in the colon. We also explored the underlying molecular pathways and networks through which these genes may contribute to disease development and progression. This study offers an excellent opportunity to investigate the role of novel and previously identified differentially regulated genes in UC. Moreover, cross-species high-resolution UC transcriptome comparison suggests new biomarkers and therapeutic targets and improves molecular phenotyping.

Methods

Mouse model, RNA sequencing, data processing, and differential expression analysis

Ulcerative colitis was induced using dextran sulfate sodium salt (DSS) in C57BL/6 male mice. The control group was time-matched and received the same drinking water without DSS. The disease activity index (DAI) was recorded. RNA was extracted from the colon and whole blood, and total and small RNA were sequenced using the Illumina HiSeq 4000 system. Gene-level quantification corresponding to the total RNA-Seq data was obtained using Stringtie (Pertea et al., 2016), and mature

miRNA expression was quantified using miRDeep2 (Friedlander et al., 2012). Differential expression was performed using DESeq2 (Love et al., 2014) (Supplementary Methods). The raw total and small RNA-Seq data have been deposited in the Gene Expression Omnibus (GEO) database with accession number GSE155303.

Functional enrichment and network analysis of differentially expressed genes

Functional enrichment analysis on the significantly differentially expressed (SDE) genes (FC > 2, padj ≤ 0.05) was performed using stringApp (Doncheva et al., 2019) in Cytoscape (Shannon et al., 2003) as well as Ingenuity pathway analysis (IPA) software (Krämer et al., 2014).

To suggest a functional role for SDE lncRNAs, we performed enrichment on their genomic neighbor genes situated within a span of 100 kb upstream and downstream of the lncRNA. For SDE miRNAs (padj ≤ 0.05), we retrieved a set of miRNA-target genes and performed enrichment on them (Supplementary Methods).

StringApp was used to retrieve STRING networks (Szklarczyk et al., 2019) in Cytoscape for the set of SDE PCGs in mouse colon and blood and for the SDE PCGs conserved in humans. Furthermore, weighted centrality analysis was performed on their STRING network to rank the SDE genes common between mouse and human colon and blood based on their importance (Supplementary Methods).

Comparison between mouse and human data

To compare our mouse data with data from UC patients, we obtained and processed data from eight publicly available studies (Van der Goten et al., 2014; Mirza et al., 2015; Vanhove et al., 2015; Lin et al., 2016; Mo et al., 2018; Haberman et al., 2019; Ostrowski et al., 2019). A set of SDE genes in the colon and blood was obtained by combining the different datasets, requiring each gene to be consistently SDE (padj ≤ 0.05) in at least two datasets. For the small RNA-Seq, we only used two datasets and required that each miRNA is SDE (padj ≤ 0.05) in at least one of these datasets. The final SDE human lists were intersected with the mouse SDE blood and colon gene sets (padj ≤ 0.05) using orthology relationships extracted from Ensembl (release 97) (Yates et al., 2020) and HCOP database (Wright et al., 2005; Seal et al., 2011) for lncRNAs (Supplementary Methods).

Validation of SDE genes by quantitative real-time PCR

The expression of 19 selected SDE genes common in mouse and human colon and blood was validated by quantitative real-time PCR (qRT-PCR). Colon and blood cDNAs from 5 DSS-UC mice, five control mice, 10 UC patients, and six controls were used for validation. Detailed information on patient demographics and primer sequences used for qRT-PCR can be found in Supplementary File S1, and procedure detail in the supplementary section. The expression of each gene tested was represented as a FC using the $2^{-\Delta\Delta CT}$ or 2 (-Delta Delta C(T)) method (Livak and Schmittgen, 2001). $GUSB$, $B2M$, $ACTB$, and TBP were the reference genes for normalization.

Results

Transcriptional profile of UC mouse model

Successful colitis induction in mice was confirmed by combined DAI score, measuring inflammatory markers, histological and micro positron emission tomography imaging analysis (Supplementary Figure S1). For the total RNA-Seq, 40–66M paired-end reads were obtained per sample, with 39–58M in the colon and 17–33M in blood, uniquely mapping to the mouse genome. For the small RNA-Seq after cleaning, 15–25M reads were obtained, where 14–20M reads mapped to the set of mouse miRNAs annotated in miRBase v.22. Principal component analysis (Supplementary Figure S2) shows a clear separation of the disease and control groups on the first principal component in total RNA-Seq colon and blood, as well as small RNA-Seq colon. The differential gene expression pattern strongly supports the separation between the groups, suggesting that global gene expression differences can be detected between UC and healthy controls (Table 1 and Supplementary File S2).

UC mouse colon and blood transcriptomics signature

We identified 3,061 and 623 SDE genes from the total RNA-Seq and 73 and 10 SDE miRNAs from the small RNA-Seq data comparing UC mice colon and blood with control, respectively (Table 1). While SDE PCGs are mainly upregulated in both tissues, lncRNAs in the colon are equally up- and downregulated, and the miRNAs in the blood are mostly downregulated. The RNA expression profile of the SDE genes and miRNAs, obtained by hierarchical clustering, is shown in heatmaps for the colon and blood (Figures 1A, C). In addition, the SDE genes and miRNAs are highlighted in the volcano plots for the colon and blood (Figures 1B, D).

From the SDE genes in the colon, we detected 116 genes with a human ortholog situated within 100 kb distance of the known Informatory bowel disease (IBD)-risk loci (Jostins et al., 2012; de Lange et al., 2017), i.e., 110 PCGs, two lncRNAs, and four other RNAs. In blood, 30 SDE genes, all PCGs, were IBD-risk loci associated (Supplementary File S3). In addition, several of the SDE genes, including $Reg3b$, $Abca12$, $Sptssb$, $Prss22$, $Pitx2$, $Ceacam12$, $Myot$, $Defb37$, Gml, and $Ugt2b5$ in the colon and $Gpr27$, $Gm7457$, $Usp2$, $Stfa3$, $Gm37800$, $Dync2li1$, $Gm32486$,

TABLE 1 Numbers of SDE genes identified in colon and blood of DSS-UC mice versus control. (For PCGs and lncRNAs |log2FC| > 1, padj ≤ 0.05 and for miRNAs FC > 1.5, padj ≤ 0.05).

	Colon	Blood
PCGs	2479 (▲1897, ▼582)	518 (▲500, ▼18)
lncRNAs	282 (▲135, ▼147)	44 (▲40, ▼4)
miRNAs	73 (▲39, ▼34)	10 (▲3, ▼7)
*Others	300 (▲201, ▼99)	54 (▲52, ▼2)

*e.g. Pseudogenes, TEC, snoRNA, miscRNA, etc. Upregulation ▲, Downregulation ▼.

FIGURE 1

DSS-UC mouse transcriptional signature. Heatmaps of SDE PCGs, lncRNAs, and miRNAs, based on z-scores of normalized log counts for **(A)** colon and **(C)** blood. Volcano plots log2FC ratio on the x-axis versus -log10 padj on the y axis for colon **(B)** and blood **(D)**. (UC, $n = 3$ and CO, $n = 3$).
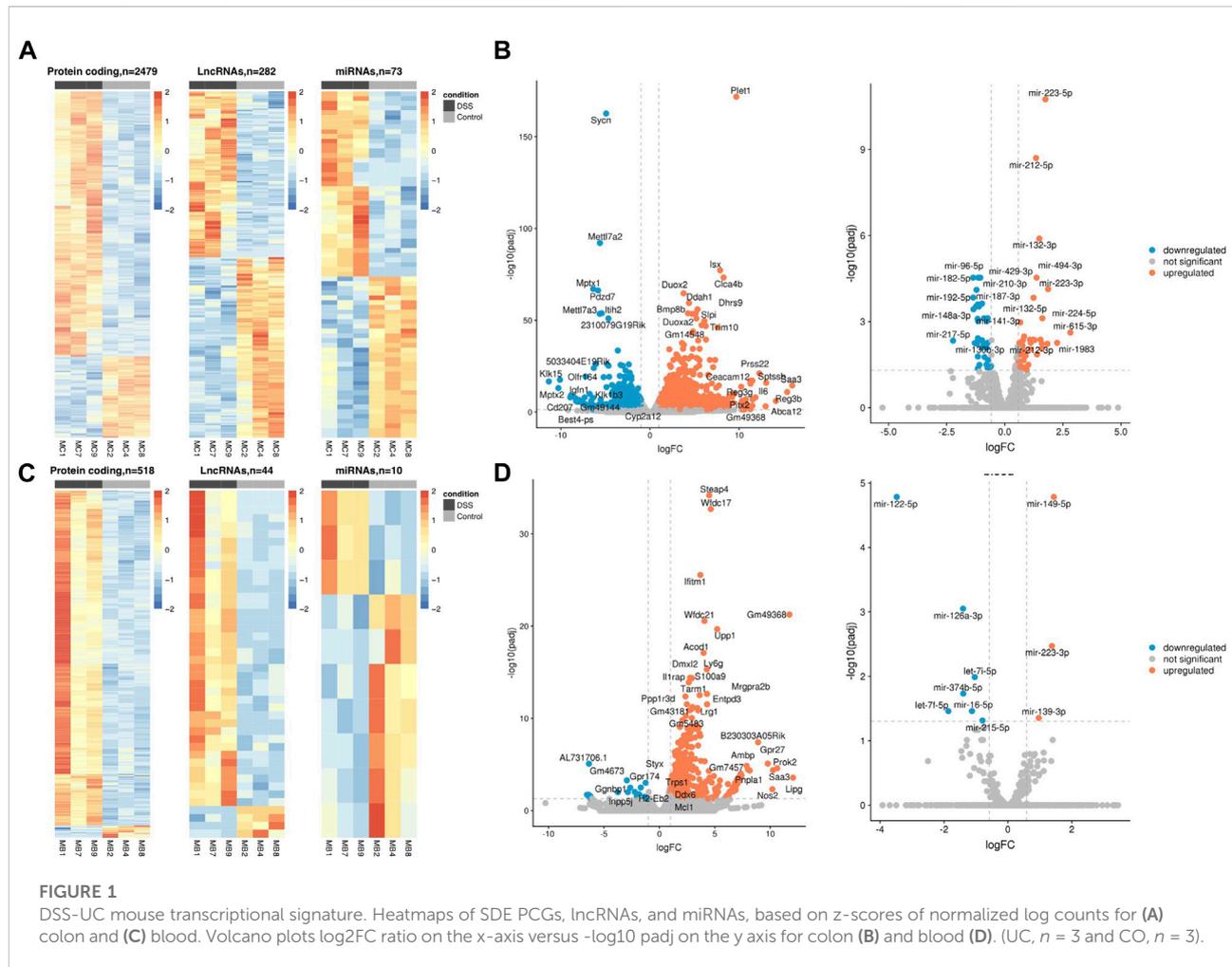
Gm7206, Gm33326, and *Epha2* in blood have not been identified previously in this model. The top 10 up- and downregulated (padj ≤ 0.05, avg. exp > 100) genes in the colon and blood are listed in Supplementary Table S1.

To investigate the roles of the SDE PCGs in the UC colon and blood, we performed a functional enrichment analysis (Supplementary File S4). The top-ranked processes and pathways for the mouse colon include the immune and inflammatory system, response to inflammation, and

organ/tissue remodeling (Supplementary Figure S3A). In particular, the KEGG disease pathway for IBD is enriched, with almost half of the pathway genes being SDE in our data. Although the number of SDE genes in blood is lower than in the colon, the top-ranked processes and pathways enriched for the SDE genes in blood overlap highly with the colon (Supplementary Figure S3B) and are mainly related to the immune system, inflammatory system, and connective tissue.

FIGURE 2
Commonly dysregulated SDE genes in UC mouse colon and blood. **(A)** Venn diagram of the SDE genes overlap between colon and blood; Other RNAs include pseudogenes, TEC, snoRNA, miscRNA, etc. Contra-regulated genes are commonly SDE genes in both tissues, which are regulated in opposite directions **(B)** MA Plot showing the relationships between $\log_2$FC in colon and blood for each of the common SDE genes; y-axis indicates whether the gene has a higher FC in colon or blood, x-axis reflects the cumulative FCs magnitude in colon and blood. Genes labeled as an outlier in each tissue were removed from the plot. **(C)** Top 15 enriched KEGG pathways and **(D)** Top 15 enriched GO biological processes for the common SDE genes between mouse blood and colon. Gene ratio corresponds to the ratio between the number of common SDE and all expressed genes in the pathway. The dot size indicates the number of SDE genes (gene counts), while the color represents the enrichment significance as given by the FDR adjusted p-values.

For the SDE lncRNAs, we performed functional enrichment of their genomic neighbors expressed in our samples (Supplementary Methods). We retrieved 888 PCGs in the colon and 159 PCGs in the blood (Supplementary File S5). Cancer, organismal injury and abnormalities, gastrointestinal disease, and immunological disease were overrepresented for both tissues, with more substantial significance in the colon than in blood (Supplementary File S5).

In the colon, for 59 out of 73 SDE miRNAs, we identified 2,330 target candidates expressed in our samples, and for blood, for eight out of 10 SDE miRNAs, we identified 1,641 targets expressed in our samples (Supplementary Figures S4A, B). In addition, functional enrichment of the targets in the colon and blood showed an overrepresentation of immune-related processes, several infectious and immunological diseases, and JAK-STAT, thyroid hormone, and MAPK signaling pathways (Supplementary File S6).

Overlap of SDE genes in the colon and blood of the UC mouse

To identify genes commonly differentially regulated in the colon and blood of UC mice with importance in the disease phenotype, 284 genes were retrieved (Supplementary File S7). Venn diagrams and an MA plot illustrating these genes' relationship in the colon and blood are shown in Figures 2A, B. From SDE common PCGs, *Prok2* (chemoattractant, labeled as an outlier), *Saa3* (acute phase apolipoproteins), and *Gm49368* (predicted gene) were the three most upregulated genes in the colon (based on logFC), while in blood, *Gm49368* was the highest followed by *Prok2* and *Saa3*. All five common lncRNAs showed significant upregulation, with *Gm11714* most upregulated in the blood and *Mirt2* most upregulated in the colon (and second-most upregulated in the blood). For the miRNAs, miR-223-3p is the only common SDE (padj ≤ 0.05), showing upregulation in both colon and blood with higher expression in blood than in the colon, possibly indicating an immune-cell enriched miRNA. However, with a less stringent threshold of padj < 0.1, we identify three other common miRNAs downregulated in both colon and blood: miR-194-5p, miR-196b-5p, and miR-215-5p. All three are highly abundant in the colon and less abundant in blood and have previously been shown to be colon tissue-specific in humans, according to the miRNA TissueAtlas database (Ludwig et al., 2016). miR-223-3p showed to be more abundant in blood compared to the others (Ludwig et al., 2016).

TABLE 2 Public human UC datasets used in our study.

Accession	Description	Platform	Method	Number of samples
Colon Total RNA-Seq				
GSE109142	Pediatric cohort, treatment-naive	RNA-Seq	Illumina TruSeq mRNA-Seq, 75 paired-End	206 UC, 20 CO
GSE117993	Pediatric cohort, treatment-naive	RNA-Seq	Illumina TruSeq mRNA-Seq, 75 Single-End	43 UC, 55 CO
GSE59071	Adult cohort	Microarray	Affymetrix GeneChip Human Gene 1.0 ST array	74 UC, 11 CO (colon)
GSE67106	Adult cohort	Microarray	Agilent Custom 8 × 60K format lncRNA expression microarray	15 UC, 9 CO (colon, rectum)
Colon Small RNA-Seq				
GSE89667	Adult cohort	Small RNA-Seq	Illumina HiSeq 2500	10 UC, 18 CO (diverticular disease)
GSE48957	Adult cohort	Microarray	Affymetrix Multispecies miRNA-2 Array	10 UC, 10 CO
Blood Total RNA-Seq				
GSE112057	Pediatric cohort	RNA-Seq	Illumina HiSeq 2000, 100 paired-End	15 UC, 12 CO
PRJEB28822-C	Pediatric cohort	RNA-Seq	Ion AmpliSeq Transcriptome	34 UC, 35 CO
PRJEB28822-A	Adult cohort	RNA-Seq	Ion AmpliSeq Transcriptome	37 UC, 32 CO

Collectively, the common SDE genes define distinct inflammatory, immune system, and connective tissue-related gene expression signatures. We observed the enrichment of genes in processes related to the immune system, specifically the regulation of inflammatory response and leukocyte activation and cytokine-cytokine receptor interactions (Figures 2C, D, Supplementary Figure S3C). In addition, several KEGG pathways are related to infectious diseases (bacterial, viral, and parasitic) that alter the immune response of cells. Analysis using IPA confirms these findings (Supplementary File S7).

Comparison between SDE genes in UC mouse colon and blood with human UC

Colon and blood transcriptomics signature of human UC

To obtain reliable sets of SDE genes in human UC and compare them with the mouse in both colon and blood, we identified and combined SDE genes from several public human datasets. For the colon, 2 RNA-Seq (Haberman et al., 2019) and 2 microarray studies (Mirza et al., 2015; Vanhove et al., 2015) and for blood, 2 RNA-Seq studies (Mo et al., 2018; Ostrowski et al., 2019) were used (Table 2, Supplementary File S8). Human miRNA data for the colon was obtained by combining one RNA-Seq and one microarray dataset. Unfortunately, we have

not been able to identify any high-quality miRNA dataset corresponding to the whole blood of UC patients.

The combined sets of human SDE PCGs and lncRNAs contained ~11,000 genes in the colon and ~2,000 genes in the blood (padj ≤ 0.05 in at least two datasets and consistently differentially regulated across datasets). We observed approximately equal proportions of up- and downregulated PCGs in the colon and a predominant (~75%) upregulation in blood. However, in both colon and blood, lncRNAs were mainly (~60%) downregulated. Furthermore, ~2,500 genes in the colon and ~130 genes in the blood were inconsistently differentially regulated between the datasets and thus not included in the combined SDE gene sets. By combining the two public small RNA-Seq datasets in the colon (Van der Goten et al., 2014; Lin et al., 2016), we obtained 207 SDE miRNAs with padj ≤ 0.05 in at least one dataset, and 37 were SDE in both datasets. No miRNA was found to be inconsistently differentially regulated between the two datasets (Table 3).

Comparison between SDE genes of human and mouse UC

Using one-to-one and one-to-many orthology assignments, we compared the combined sets of SDE genes in humans with the sets of SDE genes in mouse colon and blood (Table 4). For the colon, almost 75% of the SDE mouse genes with human orthologs are also found in the human SDE set. For blood, only 30% of the SDE mouse genes with human orthologs are

TABLE 3 Numbers of SDE genes in human colon and blood combined datasets. (For all padj ≤ 0.05 in min 2 datasets).

	Colon	Blood
PCGs	▲4658 ▼4909 ✗1531	▲1382 ▼511 ✗177
lncRNAs	▲454 ▼673 ✗812	▲12 ▼16 ✗6
miRNAs	▲14 ▼19	NA
*Others	▲450 ▼181 ✗204	▲16 ▼16 ✗6

*e.g. Pseudogenes, TEC, snoRNA, miscRNA, etc. Upregulation ▲, Downregulation ▼, Inconsistent ✗.

TABLE 4 Numbers of overlapping genes between UC mouse and UC human colon and blood (For all padj ≤ 0.05). ⇈: Upregulation in both mouse and human, ⇊: Downregulation in both mouse and human, ⇅: Opposite directions between mouse and human

	Colon			Blood		
Human SDE genes with mouse orthologs	8446 (+ 116 miRNAs)			1723		
Mouse SDE genes with human orthologs	3066 (+ 70 miRNAs)			512		
Common SDE gens in Human and Mouse	1937 (⇈1082, ⇊360, ⇅495)			159 (⇈154, ⇊3, ⇅2)		
	⇈	⇊	⇅	⇈	⇊	⇅
PCGs	1061	345	489	154	3	2
lncRNAs	5	2	1	0	0	0
miRNAs	9	12	4	NA	NA	NA
*Others	7	1	1	0	0	0

*(e.g. Pseudogenes, TEC, snoRNA, miscRNA, etc).

found in the human blood SDE gene set. The enrichment analysis for these common SDE PCGs in the UC colon and blood showed mainly immune, inflammatory, and connective tissue processes (Supplementary Files S9, 10). Seven lncRNAs showed significant up- (*H19, DNM3OS, 5430425K12RIK, 3110039I08RIK,* and *FENDRR*) and down-regulation (*HOXA11OS* and *HOTTIP*) in the colon of both mice and humans, while no common lncRNAs were detected in the blood. For five out of seven common colon SDE lncRNAs, 33 neighboring genes (100 kb upstream and downstream) were retrieved, and they were enriched for terms including connective tissue disorders, organismal injury, and abnormalities. Of the common miRNAs, twelve were downregulated, and nine were upregulated in the same direction in mouse and human colon. Functional enrichment analysis for the 1,468 targets (detected in our samples) of 17 out of 21 SDE miRNAs in the colon showed enrichment for terms including cancer, organismal injury and abnormalities, gastrointestinal disease, and several immune-related processes (Supplementary File S9).

Though ~25% of common SDE genes in the colon are contra-regulated, 356 genes with upregulation in mice and downregulation in humans showed to be mainly involved in lipid metabolism, ion transport, regulation of localization, and trans-synaptic signaling. This could reflect species-specific differences in biological processes' gene regulation. No enrichment was found for the 133 genes,

downregulated in mice and upregulated in humans. *Dio3os* was the only lncRNA in the colon with upregulation in UC mice and downregulation in humans. miR-10a-5p, miR-10b-3p, and miR-10b-5p were colon miRNAs upregulated in mice and downregulated in humans and highly expressed in both. In contrast, miR-130-3p was downregulated in mice and upregulated in humans, though lower expressed in both organisms compared to the other three miRNAs (Supplementary File S9).

Overall, 51 genes were identified as commonly SDE between the colon and blood of humans and mice compared with the healthy controls (Figure 3A). These genes were all upregulated except for *PP1R3B,* which is upregulated in blood and downregulated in the colon. From these genes, only *SLC11A1* and *STAT3* were IBD risk loci associated. Collectively, these strictly filtered genes showed to be mainly involved in inflammatory, immunological, and connective tissue-related processes (Figure 3B, Supplementary File S11). Neither of five lncRNAs and one miRNA, which were commonly SDE between the colon and blood of mice, were present in the human datasets.

From the 51 genes common between humans and mice in both colon and blood, 19 were selected for qPCR validation in a separate human UC cohort and the UC mice. Selected genes were either novel in a UC phenotype setting or not previously reported as therapeutic targets or biomarkers. IPA and Pharos were used
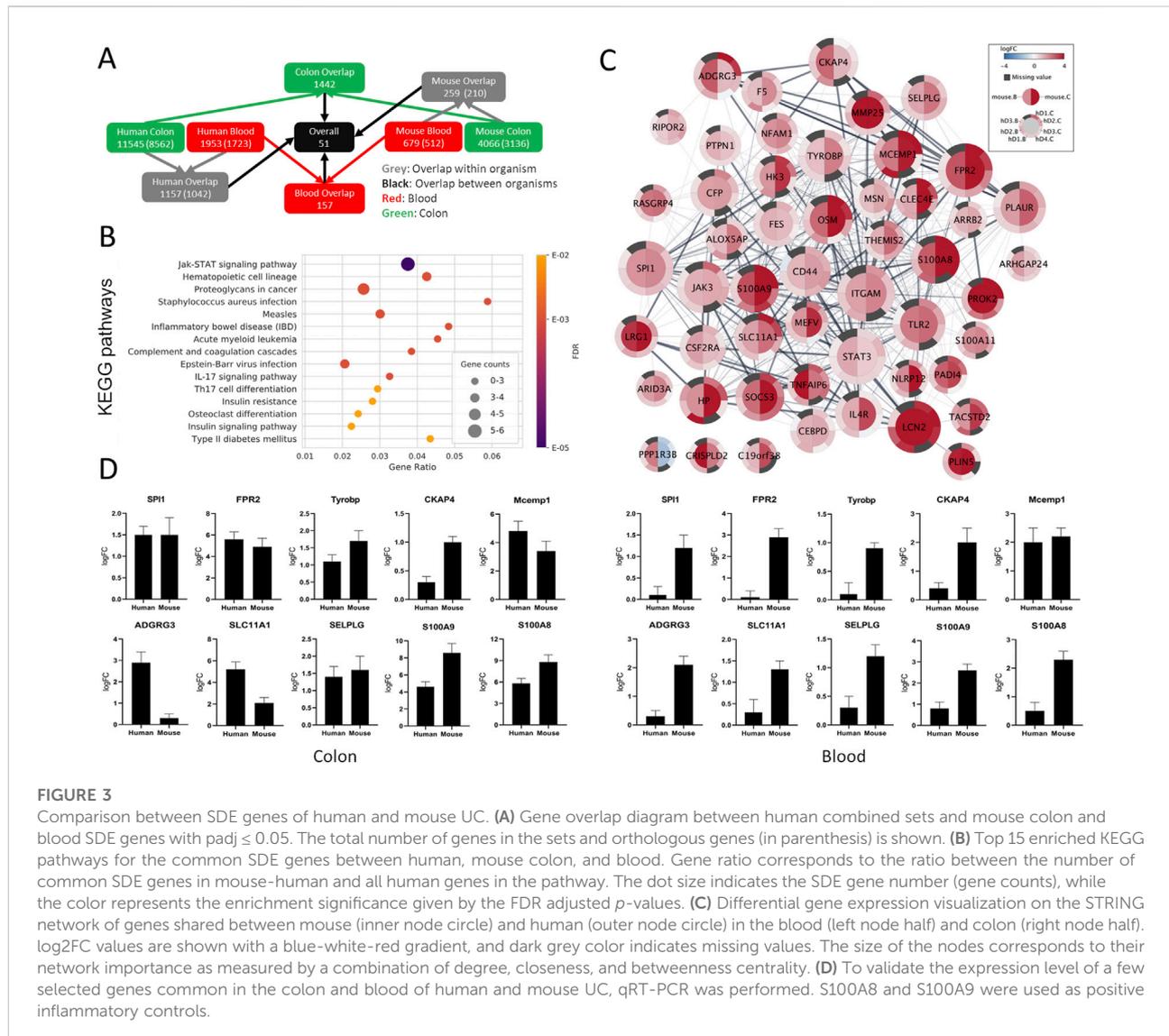
FIGURE 3
Comparison between SDE genes of human and mouse UC. **(A)** Gene overlap diagram between human combined sets and mouse colon and blood SDE genes with padj ≤ 0.05. The total number of genes in the sets and orthologous genes (in parenthesis) is shown. **(B)** Top 15 enriched KEGG pathways for the common SDE genes between human, mouse colon, and blood. Gene ratio corresponds to the ratio between the number of common SDE genes in mouse-human and all human genes in the pathway. The dot size indicates the SDE gene number (gene counts), while the color represents the enrichment significance given by the FDR adjusted *p*-values. **(C)** Differential gene expression visualization on the STRING network of genes shared between mouse (inner node circle) and human (outer node circle) in the blood (left node half) and colon (right node half). log2FC values are shown with a blue-white-red gradient, and dark grey color indicates missing values. The size of the nodes corresponds to their network importance as measured by a combination of degree, closeness, and betweenness centrality. **(D)** To validate the expression level of a few selected genes common in the colon and blood of human and mouse UC, qRT-PCR was performed. S100A8 and S100A9 were used as positive inflammatory controls.

to assess the therapeutic and biomarker applications of the 51 genes (Supplementary File S11). Furthermore, the selected genes were among the top central in the network of protein-protein associations retrieved from STRING v11 (Figure 3C, see ranking scores in Supplementary File S11) as measured by a combination of three complementary centrality measures (degree, closeness, and betweenness). While degree accounts for the direct interaction partners of each node, closeness and betweenness measure how central and thus important a node is with respect to all other nodes in the network (Ranking explained in Supplementary Methods). The expression pattern of all tested genes by qPCR in the colon and blood in both mouse and human validation cohorts nicely matched the RNA-seq analysis (Figure 3D and Supplementary Figure S5). The UC phenotype functional conservation and the importance for direct protein interactions with well-known immune-related and inflammatory

genes suggest that these genes have high therapeutic-target and diagnostic application potential.

Discussion

Over the last few years, several transcriptional profiling studies have provided evidence regarding the differential regulation of PCGs, lncRNAs, and miRNAs in UC (Mirza et al., 2015; Taman et al., 2018; Yarani et al., 2018; Haberman et al., 2019). Given the complex nature of UC, an *in vivo* model approach to studying its etiology is of paramount importance. The DSS-UC mouse has been extensively used as a surrogate to human UC. However, its transcriptional landscape and its (dis-)similarity to human UC have not yet been studied and comprehensively characterized.

Here, we performed colon and blood transcriptional profiling of UC mice for coding and non-coding RNAs and investigated the affected biological pathways. Our analyses showed that most differentially regulated genes are upregulated and involved in immunological and inflammatory responses and connective tissue, organ abnormality, and injuries. This points to the pathological nature of UC and the importance of inflammatory processes in destroying epithelium and aggravation of the disease. Moreover, alterations and mainly upregulation of genes that suppress inflammation were detected in the colon and blood. Genes, which demonstrated significant downregulation, were mainly genes with basic metabolic functions involved in cellular growth and proliferation, lipid, vitamin, amino acid, and mineral metabolism. Most of the genes common between the colon and blood of UC mice are regulated in the same direction (upregulated) and are related to immune or inflammatory processes. Among these genes, *S100a8, S100a9, Il1b, Lcn2, C3, Cxcl2, Nos2, Slpi, Socs3, Ifitm6, Csf2rb, Tlr2,* and *Cd44* have previously been detected in UC mouse models in other studies (te Velde et al., 2007; Hansen et al., 2009; Fang et al., 2011; Holgersen et al., 2015; Rankin et al., 2018). Several other SDE genes in the colon and blood from our study, including *Lipg, Gm5483, B230303A05Rik, Ambp,* and *Trem3,* have not been shown previously in this model. The largest overlap of 1,566 SDE genes was with a recent study by Czarnewski et al. (2019) where deep RNA sequencing was performed to stratify different subtypes of human UC using the common genes in human and mouse UC. Although they used a very similar approach, their study did not consider non-coding RNAs nor blood transcriptomes.

Here we investigated the transcriptional landscape of both lncRNAs and miRNAs in the colon and blood of the DSS-UC mice model and identified widespread differential regulation of non-coding RNAs. Numerous SDE lncRNAs and miRNAs were identified in both tissues, many of which have not been reported previously. Several SDE lncRNAs in the colon have been widely studied previously, including *H19, Meg3, Hottip, Hoxa11os,* and *Mirt2*. The latter was also SDE in blood and is a negative regulator of inflammation (Du et al., 2017). *H19* upregulation is believed to decrease the expression of the vitamin D receptor significantly and thus have a destructive effect on the intestinal epithelial barrier function by increasing permeability and decreasing the expression of ZO-1 and occludin tight junction proteins (Chen et al., 2016; Yarani et al., 2018). Not much is known regarding the other lncRNAs in UC settings, but recently *Meg3* was shown to inhibit the inflammatory response in ankylosing spondylitis characterized by chronic inflammation (Li et al., 2020). *Hottip* silencing in rheumatoid arthritis led to reduced inflammation (Hu et al., 2020). In UC mouse blood, lncRNAs *Neat1* and *Panct2* also showed to be SDE in disease. *Neat1* is shown to be proinflammatory, and its inhibition suppresses the inflammatory response in IBD (Liu et al., 2018), whereas little is known about *Panct2*. Moreover,

numerous SDE lncRNAs were not reported before in UC mouse, including the top three upregulated *9130221F21Rik, E230034D01Rik, 1200007C13Rik* in the colon, and *Gm7457, Gm37800, Gm32486* in blood, and many more up- and downregulated new lncRNAs were identified in our study.

Among the SDE miRNAs, several IBD-related miRNAs, including miR-30, miR-223, miR-21, and miR-142, were upregulated, and miR-192 was downregulated in the colon, while miR-223, miR-16, and miR-126 were upregulated in blood. miR-223-3p is the only miR upregulated in both colon and blood of the UC mice. It is upregulated in neutrophils and monocytes and acts as a controller of NLRP3 inflammasome activity, regulating the intestine inflammatory process by affecting IL-1β production (Neudecker et al., 2017). It has also been shown that miR-223 mediates the cross-talk between the intestinal barrier and the IL-23 pathway by targeting CLDN8, a claudin protein that constitutes the backbone of the intestinal barrier (Wang et al., 2016a; Yarani et al., 2022). miR-223 has also been used as a biomarker in IBD (Wang et al., 2016b; Yarani et al., 2022). Thus, the evidence suggests its proinflammatory role and highlights its potential as an RNA biomarker that seems to be conserved between different species. It is noteworthy that in the current study, many new SDE miRNAs are identified which are not shown in the UC mouse model previously, including the top three upregulated miR-615-3p, miR-212-3p, miR-224-5p in the colon, and miR-149-5p in blood, and many more, which need future detailed functional investigations.

By comparing our UC mouse to human UC data, we identified similar widespread directional differential regulation of genes. The overlap of a substantial number of differentially regulated genes in mouse and human UC suggests common fundamental mechanisms responsible for disease onset and progression. To date, few colon studies have compared the transcriptional changes in UC mouse models and human UC (Holgersen et al., 2015; Rankin et al., 2018; Czarnewski et al., 2019). Holgersen et al. (2015) identified 92 SDE genes in both human CD and UC, which were used to compare their three mouse models. Since colon samples were pooled in the DSS model, no statistical significance of differences could be calculated between disease and healthy mice. However, 59 genes showed to be shared between our lists. In their study, non-coding RNAs and blood transcriptomes were not studied. Rankin et al. (2018) however, identified 12 PCGs and six lncRNAs associated with DSS mouse colitis and human UC. No small RNA and blood profiling were performed in their study. In the Czarnewski et al. (2019) study, 650 SDE genes common between human and mouse UC colon were identified, but no non-coding RNA and blood profiling was performed. Apart from these studies, to the best of our knowledge, there are no human-mouse UC comparative studies considering both colon and blood and coding and non-coding RNAs to the scale of our research.

Although we recognize that separate analysis of common genes in the colon and blood may provide additional

information, we focus on the genes that were common between both tissues in the mouse and human, thus, may represent the universal markers of UC. Ascribed to the pathophysiology of the UC and the fact that for diagnosis purposes, it is difficult to access the colon in comparison to blood, identification of biomarker candidates showing a similar differential expression in the colon and blood helps us diagnose UC faster and without the need for invasive assessments. Moreover, conserved same-direction differentially expressed genes in two distinct species might serve as a more reliable disease-specific biomarker compared with genes that are not expressed in one species or expressed differently in different species under the same pathogenic condition.

Overall, 51 common PCGs in mouse-human colon and blood with strong inflammatory and immunological profiles showed to be consistently differentially regulated. Several of these genes, including *ITGAM, STAT3, LCN2, TLR2, PLAUR, JAK3, CD44, OSM, SOCS3, S100A8, S100A9, HP, CFP, IL4R, FES,* and *CSF2RA,* have previously been associated with IBD. Interestingly, several less known/studied genes in IBD, including *SPI1, FPR2, TYROBP, CKAP4, MCEMP1, ADGRG3, SLC11A1,* and *SELPLG,* were identified among the top candidates. *SPI1* encodes an ETS-domain transcription factor PU.1 that activates gene expression exclusively in hematopoietic cells, including myeloid and lymphoid cells, e.g., such as lymphocyte B cell development, and is involved in disease-like Inflammatory diarrhea, primary mediastinal B cell Lymphoma and pediatric T cell acute lymphoblastic leukemia (Seki et al., 2017; Rothenberg et al., 2019). *FPR2* works as a chemoattractant receptor involved in antibacterial host defense and inflammation by sensing bacteria (Alessi et al., 2017), which is expressed not only by immune cells but also by epithelial, endothelial, and fibroblasts cells (He and Ye, 2017) that elicit proinflammatory responses. *FPR2* also promotes monocyte inflammatory activities, and its absence in knock-out mice results in increased bacteria load in the liver and reduced neutrophil infiltration (Carp, 1982). *TYROBP* encodes a transmembrane signaling polypeptide associated with the killer-cell inhibitory receptor family and plays a role in signal transduction and inflammation (Tomasello and Vivier, 2005). Loss of *TYROBP* has resulted in presenile dementia with bone cysts (Paloneva et al., 2000). *TYROBP* expression is increased in Alzheimer's disease (AD) patients, and its deficiency in AD mice showed to be neuroprotective and immune-inflammatory therapeutic, which eventually slowed/arrested the progression of pathological late-onset sporadic AD (Haure-Mirande et al., 2017). *CKAP4* (also known as *CLIMP-63*) encodes a transmembrane protein and has been the focus of several investigations (reviewed in (Tuffy and Planey, 2012)) and is believed to regulate cell migration (Osugi et al., 2019). *MCEMP1* encodes a single-pass transmembrane protein involved in immune responses through mast cell differentiation (Li et al., 2005). *ADGRG3* has been shown to regulate the antimicrobial activity of granulocytes (Hsiao et al., 2018) and seems to be required for macrophages' local inflammation development (Shi et al., 2016). *SLC11A1*, known as *NRAMP1*, is expressed exclusively in immune monocytes and

phagocytes like macrophages (Canonne-Hergaux et al., 1999). Mutation in these genes may affect susceptibility to infections and autoimmune diseases (Sechi et al., 2006). *SLC11A1* product acidifies the phagosome (Lapham et al., 2004), thereby killing entrapped pathogens. *SELPLG*'s strong association with the immune system has been emphasized before (Tinoco et al., 2016). This gene codes for *PSGL-1* protein, a counter receptor for P-selectin that facilitates immune responses by promoting immune effector-cells trafficking into inflamed tissue. These genes are evolutionary conserved and have known interactions with other already well-known IBD candidate genes. Thus, these candidates potentially represent new UC diagnostic and therapeutic targets and novel avenues for more detailed disease mechanistic studies.

No long non-coding RNAs made it to the 51 final gene list. The major reason is that only a few lncRNAs are assigned orthology relationships across species in general (e.g., only 131 of all mouse lncRNAs are assigned a human ortholog based on the HCOP database). However, we have identified seven lncRNAs that were SDE in the colon in both humans and mice. Apart from H19, little is known about the other six lncRNAs in IBD, which clearly shows a need for detailed analysis. We also identified 21 miRNAs SDE in the colon in humans and mice. Among them, miR-146a-5p, miR -155-5p, miR-192-5p, miR-194-5p, miR-196b-5p, miR-200c-3p, miR-223-3p, and miR-223-5p were frequently shown to be differentially regulated in IBD previously, which makes them a promising candidate for further analysis.

Given the complexity of the whole colon and blood cell types, one could isolate different cell populations and perform single-cell sequencing to distinguish pathogenic mechanisms with higher resolution. One of the limitations of our study is that whole colon and blood gene expression profiling was performed. Albeit, the primary rationale was to preserve the natural state of the disease as much as possible. In addition, the disadvantages of isolation of different cell types are the technical fractionation procedures and the time duration from tissue collection to sample processing, potentially affecting gene expression statuses specifically for non-coding RNAs in cells.

In conclusion, high-throughput transcriptome analysis provides a unique tool for discovering new therapeutic and diagnostic targets by identifying relationships between gene expression and disease phenotype. Here, we showed that a one-to-one comparison of the transcriptome of the DSS-UC mouse model to human UC could provide novel disease pathophysiology information by identifying genes and pathways with essential contributions to UC. Our data suggest that prospective therapeutic interventions and diagnostic applications should target multiple major gene regulators involved in UC pathogenesis and propagations and combine several genes for valid biomarker applications. Moreover, targeting genes with conserved functional roles in the disease pathogenesis may offer

a reliable UC treatment approach compared to targets with different functional roles in different tissues and organisms.

Data availability statement

The data underlying this article are available in the article and its online Supplementary Material. The raw total and small RNA-Seq data have been deposited in the Gene Expression Omnibus (GEO) database with accession number GSE155303.

Ethics statement

The studies involving human participants were reviewed and approved by the Regional Ethical Committee (H-4-2012-030). The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by Danish Animal Experiments Inspectorate, Ministry of Food, Agriculture, and Fisheries, Denmark, at Odense University Hospital under the 2015-15-0201-00730 license number.

Author contributions

RY designed the study, established the mouse model, performed the experiments, interpreted the results, and wrote the manuscript (Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Writing—Original Draft, Visualization, Project administration). OP processed and analyzed the RNA-Seq and public datasets and contributed to writing the manuscript (Software, Formal analysis, Data Curation, Writing—Original Draft, Visualization). ND contributed to analyzing the data and to writing the manuscript (Software, Formal analysis, Writing—Original Draft, Visualization). CA provided support in analyzing the data and revised the manuscript (Formal analysis). BP and UH provided tools and help for establishing the mouse model (Methodology, Resources). CS performed the qPCRs (Validation). AM prepared the human samples and revised the manuscript (Resources). TL provided essential insights and revised the manuscript (Conceptualization). CB-B and MV prepared the human samples and revised the manuscript (Resources). JG and LJ provided essential tools and insights in designing the experiment supervised the study and revised the manuscript (Conceptualization, Supervision, Funding acquisition). FP contributed to designing the experiment, provided essential tools and insights, supervised the study, and revised the manuscript (Conceptualization, Supervision, Project administration, Funding acquisition). All authors read, reviewed, and approved the final manuscript.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.1081176/full#supplementary-material

SUPPLEMENTARY FILE S1
Primers for qPCR and Patients info.

SUPPLEMENTARY FILE S2
SDEGs in colon and blood.

SUPPLEMENTARY FILE S3
List of IBD loci genes.

SUPPLEMENTARY FILE S4
Functional enrichment for SDE PCGs in colon and blood.

SUPPLEMENTARY FILE S5
lncRNA-neighbors and functional enrichment.

SUPPLEMENTARY FILE S6
miRNA-targets and functional enrichment.

SUPPLEMENTARY FILE S7
Common SDEGs in mouse blood and colon.

SUPPLEMENTARY FILE S8
Confident sets SDEGs in Human UC.

SUPPLEMENTARY FILE S9
Colon common SDEGs in mouse and human.

SUPPLEMENTARY FILE S10
Blood common SDEGs in mouse and human.

SUPPLEMENTARY FILE S11
Final 51 SDEGs, centrality ranking and applications.

SUPPLEMENTARY FILE S12
Supplementary methods and figures.

References

Alessi, M. C., Cenac, N., Si-Tahar, M., and Riteau, B. (2017). FPR2: A novel promising target for the treatment of influenza. *Front. Microbiol.* 8, 1719. doi:10.3389/fmicb.2017.01719

Canonne-Hergaux, F., Gruenheid, S., Govoni, G., and Gros, P. (1999). The Nramp1 protein and its role in resistance to infection and macrophage function. *Proc. Assoc. Am. Physicians* 111 (4), 283–289. doi:10.1046/j.1525-1381.1999.99236.x

Carp, H. (1982). Mitochondrial N-formylmethionyl proteins as chemoattractants for neutrophils. *J. Exp. Med.* 155 (1), 264–275. doi:10.1084/jem.155.1.264

Casamassimi, A., Federico, A., Rienzo, M., Esposito, S., and Ciccodicola, A. (2017). Transcriptome profiling in human diseases: New advances and perspectives. *Int. J. Mol. Sci.* 18 (8), 1652. doi:10.3390/ijms18081652

Chassaing, B., Aitken, J. D., Malleshappa, M., and Vijay-Kumar, M. (2014). Dextran sulfate sodium (DSS)-induced colitis in mice. *Curr. Protoc. Immunol.* 104, 15 25 1–15. doi:10.1002/0471142735.im1525s104

Chen, S. W., Wang, P. Y., Liu, Y. C., Sun, L., Zhu, J., Zuo, S., et al. (2016). Effect of long noncoding RNA H19 overexpression on intestinal barrier function and its potential role in the pathogenesis of ulcerative colitis. *Inflamm. Bowel Dis.* 22 (11), 2582–2592. doi:10.1097/MIB.0000000000000932

Czarnewski, P., Parigi, S. M., Sorini, C., Diaz, O. E., Das, S., Gagliani, N., et al. (2019). Conserved transcriptomic profile between mouse and human colitis allows unsupervised patient stratification. *Nat. Commun.* 10 (1), 2892. doi:10.1038/s41467-019-10769-x

de Lange, K. M., Moutsianas, L., Lee, J. C., Lamb, C. A., Luo, Y., Kennedy, N. A., et al. (2017). Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat. Genet.* 49 (2), 256–261. doi:10.1038/ng.3760

Doncheva, N. T., Morris, J. H., Gorodkin, J., and Jensen, L. J. (2019). Cytoscape StringApp: Network analysis and visualization of proteomics data. *J. Proteome Res.* 18 (2), 623–632. doi:10.1021/acs.jproteome.8b00702

Du, M., Yuan, L., Tan, X., Huang, D., Wang, X., Zheng, Z., et al. (2017). The LPS-inducible lncRNA Mirt2 is a negative regulator of inflammation. *Nat. Commun.* 8 (1), 2049. doi:10.1038/s41467-017-02229-1

Eichele, D. D., and Kharbanda, K. K. (2017). Dextran sodium sulfate colitis murine model: An indispensable tool for advancing our understanding of inflammatory bowel diseases pathogenesis. *World J. Gastroenterol.* 23 (33), 6016–6029. doi:10.3748/wjg.v23.i33.6016

Fang, K., Bruce, M., Pattillo, C. B., Zhang, S., Stone, R., Clifford, J., et al. (2011). Temporal genomewide expression profiling of DSS colitis reveals novel inflammatory and angiogenesis genes similar to ulcerative colitis. *Physiol. Genomics* 43 (1), 43–56. doi:10.1152/physiolgenomics.00138.2010

Friedlander, M. R., Mackowiak, S. D., Li, N., Chen, W., and Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* 40 (1), 37–52. doi:10.1093/nar/gkr688

Haberman, Y., Karns, R., Dexheimer, P. J., Schirmer, M., Somekh, J., Jurickova, I., et al. (2019). Ulcerative colitis mucosal transcriptomes reveal mitochondriopathy and personalized mechanisms underlying disease severity and treatment response. *Nat. Commun.* 10 (1), 38. doi:10.1038/s41467-018-07841-3

Hansen, J. J., Holt, L., and Sartor, R. B. (2009). Gene expression patterns in experimental colitis in IL-10-deficient mice. *Inflamm. Bowel Dis.* 15 (6), 890–899. doi:10.1002/ibd.20850

Hasler, R., Sheibani-Tezerji, R., Sinha, A., Barann, M., Rehman, A., Esser, D., et al. (2017). Uncoupling of mucosal gene regulation, mRNA splicing and adherent microbiota signatures in inflammatory bowel disease. *Gut* 66 (12), 2087–2097. doi:10.1136/gutjnl-2016-311651

Haure-Mirande, J. V., Audrain, M., Fanutza, T., Kim, S. H., Klein, W. L., Glabe, C., et al. (2017). Deficiency of TYROBP, an adapter protein for TREM2 and CR3 receptors, is neuroprotective in a mouse model of early Alzheimer's pathology. *Acta Neuropathol.* 134 (5), 769–788. doi:10.1007/s00401-017-1737-3

He, H. Q., and Ye, R. D. (2017). The formyl peptide receptors: Diversity of ligands and mechanism for recognition. *Molecules* 22 (3), 455. doi:10.3390/molecules22030455

Holgersen, K., Kutlu, B., Fox, B., Serikawa, K., Lord, J., Hansen, A. K., et al. (2015). High-resolution gene expression profiling using RNA sequencing in patients with inflammatory bowel disease and in mouse models of colitis. *J. Crohns Colitis* 9 (6), 492–506. doi:10.1093/ecco-jcc/jjv050

Hong, S. N., Joung, J. G., Bae, J. S., Lee, C. S., Koo, J. S., Park, S. J., et al. (2017). RNA-Seq reveals transcriptomic differences in inflamed and noninflamed intestinal mucosa of crohn's disease patients compared with normal mucosa of healthy controls. *Inflamm. Bowel Dis.* 23 (7), 1098–1108. doi:10.1097/MIB.0000000000001066

Hsiao, C. C., Chu, T. Y., Wu, C. J., van den Biggelaar, M., Pabst, C., Hebert, J., et al. (2018). The adhesion G protein-coupled receptor GPR97/ADGRG3 is expressed in human granulocytes and triggers antimicrobial effector functions. *Front. Immunol.* 9, 2830. doi:10.3389/fimmu.2018.02830

Hu, X., Tang, J., Bao, P., Deng, W., Wu, J., Liang, Y., et al. (2020). Silencing of long non-coding RNA HOTTIP reduces inflammation in rheumatoid arthritis by demethylation of SFRP1. *Mol. Ther. Nucleic Acids* 19, 468–481. doi:10.1016/j.omtn.2019.11.015

Jostins, L., Ripke, S., Weersma, R. K., Duerr, R. H., McGovern, D. P., Hui, K. Y., et al. (2012). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 491 (7422), 119–124. doi:10.1038/nature11582

Krämer, A., Green, J., Pollard, J., and Tugendreich, S. (2014). Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* 30 (4), 523–530. doi:10.1093/bioinformatics/btt703

Lapham, A. S., Phillips, E. S., and Barton, C. H. (2004). Transcriptional control of Nramp1: A paradigm for the repressive action of c-myc. *Biochem. Soc. Trans.* 32 (6), 1084–1086. doi:10.1042/BST0321084

Lekka, E., and Hall, J. (2018). Noncoding RNAs in disease. *FEBS Lett.* 592 (17), 2884–2900. doi:10.1002/1873-3468.13182

Li, K., Wang, S. W., Li, Y., Martin, R. E., Li, L., Lu, M., et al. (2005). Identification and expression of a new type II transmembrane protein in human mast cells. *Genomics* 86 (1), 68–75. doi:10.1016/j.ygeno.2005.03.006

Li, Y., Zhang, S., Zhang, C., and Wang, M. (2020). LncRNA MEG3 inhibits the inflammatory response of ankylosing spondylitis by targeting miR-146a. *Mol. Cell Biochem.* 466 (1-2), 17–24. doi:10.1007/s11010-019-03681-x

Lin, J., Zhang, X., Zhao, Z., Welker, N. C., Li, Y., Liu, Y., et al. (2016). Novel MicroRNA signature to differentiate ulcerative colitis from crohn disease: A genome-wide study using next generation sequencing. *Microrna* 5 (3), 222–229. doi:10.2174/2211536605666161117113031

Liu, R., Tang, A., Wang, X., Chen, X., Zhao, L., Xiao, Z., et al. (2018). Inhibition of lncRNA NEAT1 suppresses the inflammatory response in IBD by modulating the intestinal epithelial barrier and by exosome-mediated polarization of macrophages. *Int. J. Mol. Med.* 42 (5), 2903–2913. doi:10.3892/ijmm.2018.3829

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25 (4), 402–408. doi:10.1006/meth.2001.1262

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15 (12), 550. doi:10.1186/s13059-014-0550-8

Ludwig, N., Leidinger, P., Becker, K., Backes, C., Fehlmann, T., Pallasch, C., et al. (2016). Distribution of miRNA expression across human tissues. *Nucleic Acids Res.* 44 (8), 3865–3877. doi:10.1093/nar/gkw116

Mirza, A. H., Berthelsen, C. H., Seemann, S. E., Pan, X., Frederiksen, K. S., Vilien, M., et al. (2015). Transcriptomic landscape of lncRNAs in inflammatory bowel disease. *Genome Med.* 7 (1), 39. doi:10.1186/s13073-015-0162-2

Mo, A., Marigorta, U. M., Arafat, D., Chan, L. H. K., Ponder, L., Jang, S. R., et al. (2018). Disease-specific regulation of gene expression in a comparative analysis of juvenile idiopathic arthritis and inflammatory bowel disease. *Genome Med.* 10 (1), 48. doi:10.1186/s13073-018-0558-x

Neudecker, V., Haneklaus, M., Jensen, O., Khailova, L., Masterson, J. C., Tye, H., et al. (2017). Myeloid-derived miR-223 regulates intestinal inflammation via repression of the NLRP3 inflammasome. *J. Exp. Med.* 214 (6), 1737–1752. doi:10.1084/jem.20160462

Okayasu, I., Hatakeyama, S., YaMadaM.Ohkusa, T., Inagaki, Y., and Nakaya, R. (1990). A novel method in the induction of reliable experimental acute and chronic ulcerative colitis in mice. *Gastroenterology* 98 (3), 694–702. doi:10.1016/0016-5085(90)90290-h

Ostrowski, J., Dabrowska, M., Lazowska, I., Paziewska, A., Balabas, A., Kluska, A., et al. (2019). Redefining the practical utility of blood transcriptome biomarkers in inflammatory bowel diseases. *J. Crohns Colitis* 13 (5), 626–633. doi:10.1093/ecco-jcc/jjy205

Osugi, Y., Fumoto, K., and Kikuchi, A. (2019). CKAP4 regulates cell migration via the interaction with and recycling of integrin. *Mol. Cell Biol.* 39 (16), e00073. doi:10.1128/MCB.00073-19

Paloneva, J., KestilaM.Wu, J., Salminen, A., Bohling, T., Ruotsalainen, V., et al. (2000). Loss-of-function mutations in TYROBP (DAP12) result in a presenile dementia with bone cysts. *Nat. Genet.* 25 (3), 357–361. doi:10.1038/77153

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11 (9), 1650–1667. doi:10.1038/nprot.2016.095

Rankin, C. R., Theodorou, E., Man Law, I. K., Rowe, L., Kokkotou, E., Pekow, J., et al. (2018). Identification of novel mRNAs and lncRNAs associated with mouse experimental colitis and human inflammatory bowel disease. *Am. J. Physiol. Gastrointest. Liver Physiol.* 315 (5), G722–G733. doi:10.1152/ajpgi.00077.2018

Rothenberg, E. V., Hosokawa, H., and Ungerback, J. (2019). Mechanisms of action of hematopoietic transcription factor PU.1 in initiation of T-cell development. *Front. Immunol.* 10, 228. doi:10.3389/fimmu.2019.00228

Seal, R. L., Gordon, S. M., Lush, M. J., Wright, M. W., and Bruford, E. A. (2011). genenames.org: the HGNC resources in 2011. *Nucleic Acids Res.* 39, D514–D519. doi:10.1093/nar/gkq892

Sechi, L. A., Gazouli, M., Sieswerda, L. E., Molicotti, P., Ahmed, N., Ikonomopoulos, J., et al. (2006). Relationship between Crohn's disease, infection with *Mycobacterium avium* subspecies paratuberculosis and SLC11A1 gene polymorphisms in Sardinian patients. *World J. Gastroenterol.* 12 (44), 7161–7164. doi:10.3748/wjg.v12.i44.7161

Seki, M., Kimura, S., Isobe, T., Yoshida, K., Ueno, H., Nakajima-Takagi, Y., et al. (2017). Recurrent SPI1 (PU.1) fusions in high-risk pediatric T cell acute lymphoblastic leukemia. *Nat. Genet.* 49 (8), 1274–1281. doi:10.1038/ng.3900

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303

Shi, J., Zhang, X., Wang, S., Wang, J., Du, B., Wang, Z., et al. (2016). Gpr97 is dispensable for metabolic syndrome but is involved in macrophage inflammation in high-fat diet-induced obesity in mice. *Sci. Rep.* 6, 24649. doi:10.1038/srep24649

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47 (D1), D607–D613. doi:10.1093/nar/gky1131

Taman, H., Fenton, C. G., Hensel, I. V., Anderssen, E., Florholmen, J., and Paulssen, R. H. (2018). Transcriptomic landscape of treatment-naïve ulcerative colitis. *J. Crohns Colitis* 12 (3), 327–336. doi:10.1093/ecco-jcc/jjx139

te Velde, A. A., de Kort, F., Sterrenburg, E., Pronk, I., ten Kate, F. J. W., Hommes, D. W., et al. (2007). Comparative analysis of colonic gene expression of three experimental colitis models mimicking inflammatory bowel disease. *Inflamm. Bowel Dis.* 13 (3), 325–330. doi:10.1002/ibd.20079

Tinoco, R., Carrette, F., Barraza, M. L., Otero, D. C., Magana, J., Bosenberg, M. W., et al. (2016). PSGL-1 is an immune checkpoint regulator that promotes T cell exhaustion. *Immunity* 44 (5), 1190–1203. doi:10.1016/j.immuni.2016.04.015

Tomasello, E., and Vivier, E. (2005). KARAP/DAP12/TYROBP: Three names and a multiplicity of biological functions. *Eur. J. Immunol.* 35 (6), 1670–1677. doi:10.1002/eji.200425932

Tuffy, K. M., and Planey, S. L. (2012). Cytoskeleton-associated protein 4: Functions beyond the endoplasmic reticulum in Physiology and disease. *ISRN Cell Biol.* 2012, 1–11. doi:10.5402/2012/142313

Van der Goten, J., Vanhove, W., Lemaire, K., Van Lommel, L., Machiels, K., Wollants, W. J., et al. (2014). Integrated miRNA and mRNA expression profiling in inflamed colon of patients with ulcerative colitis. *PLoS One* 9 (12), e116117. doi:10.1371/journal.pone.0116117

Vanhove, W., Peeters, P. M., Staelens, D., Schraenen, A., Van der Goten, J., Cleynen, I., et al. (2015). Strong upregulation of AIM2 and IFI16 inflammasomes in the mucosa of patients with active inflammatory bowel disease. *Inflamm. Bowel Dis.* 21 (11), 2673–2682. doi:10.1097/MIB.0000000000000535

Wang, H., Chao, K., Ng, S. C., Bai, A. H., Yu, Q., Yu, J., et al. (2016). Pro-inflammatory miR-223 mediates the cross-talk between the IL23 pathway and the intestinal barrier in inflammatory bowel disease. *Genome Biol.* 17, 58. doi:10.1186/s13059-016-0901-8

Wang, H., Zhang, S., Yu, Q., Yang, G., Guo, J., Li, M., et al. (2016). Circulating MicroRNA223 is a new biomarker for inflammatory bowel disease. *Med. Baltim.* 95 (5), e2703. doi:10.1097/MD.0000000000002703

Wright, M. W., Eyre, T. A., Lush, M. J., Povey, S., and Bruford, E. A. (2005). HCOP: The HGNC comparison of orthology predictions search tool. *Mamm. Genome* 16 (11), 827–828. doi:10.1007/s00335-005-0103-2

Yarani, R., Mirza, A. H., Kaur, S., and Pociot, F. (2018). The emerging role of lncRNAs in inflammatory bowel disease. *Exp. Mol. Med.* 50 (12), 1–14. doi:10.1038/s12276-018-0188-9

Yarani, R., Shojaeian, A., Palasca, O., Doncheva, N. T., Jensen, L. J., Gorodkin, J., et al. (2022). Differentially expressed miRNAs in ulcerative colitis and crohn's disease. *Front. Immunol.* 13, 865777. doi:10.3389/fimmu.2022.865777

Yates, A. D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., et al. (2020). Ensembl 2020. *Nucleic Acids Res.* 48 (D1), D682–D688. doi:10.1093/nar/gkz966