Integrated multi-model analysis of intestinal inflammation exposes key molecular features of preclinical and clinical IBD

Miguel González-Acera[1,#], Jay V. Patankar[1,#], Lena Erkert[1], Roodline Cineus[2,3], Reyes Gámez-Belmonte[1], Tamara Leupold[1], Marvin Bubeck[1], Li-Li Bao[1], Martin Dinkel[1], Ru Wang[1], Heidi Limberger[1], Iris Stolzer[1], Katharina Gerlach[1], Fabrizio Mascia[1], Kristina Koop[1], Christina Plattner[6], Gregor Sturm[6], Benno Weigmann[1], Claudia Günther[1], Stefan Wirtz[1], Kai Hildner[1,5], Anja A. Kühl[2,4], Britta Siegmund[2], Raja Atreya[1,5], Markus F. Neurath[1,5], The IBDome Consortium[7], Ahmed N. Hegazy[2,3], Zlatko Trajanoski[6], Christoph Becker[1,5 *]

1 Department of Medicine 1, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) and Universitätsklinikum Erlangen, Erlangen, Germany.

2 Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Medizinische Klinik für Gastroenterologie, Infektiologie und Rheumatologie, Berlin, Germany

3 Deutsches Rheuma-Forschungszentrum, ein Institut der Leibniz-Gemeinschaft, 10117 Berlin, Germany

4 Charité - Universitätsmedizin Berlin, Freie Universität Berlin, Humboldt Universität zu Berlin and Berlin Institute of Health, iPATH.Berlin, Berlin, Germany

5 Deutsches Zentrum Immuntherapie (DZI), Erlangen, Germany

6 Biocenter, Institute of Bioinformatics, Medical University of Innsbruck, Innsbruck, Austria.

7 The IBDome Consortium, see under other contributing authors

[#] Share first-authorship

* Corresponding author

Keywords: Inflammatory bowel disease, intestinal epithelial cells, mitochondrial dynamics, microvilli, comparative transcriptomics

Short title: Refining preclinical modeling of IBD

Address correspondence to:

Christoph Becker, PhD

Department of Medicine 1, Universitätsklinikum Erlangen

Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Erlangen, Germany

Phone: (0049) 9131-85 35 886

Email: christoph.becker@uk-erlangen.de

Abstract

Inflammatory bowel disease (IBD) is a chronic inflammatory condition of the intestine with a complex and multifaceted pathogenesis. While various animal models exist to study specific disease mechanisms relevant to human IBD, a comprehensive comparative framework linking these to IBD pathophysiology is lacking. In this study, we provide a framework that delineates common and unique features encountered at the transcriptomic level in 13 widely used mouse models, employing both curation-based and statistically correlative analyses. Our comparative transcriptomic analyses between mouse models versus established as well as new patient datasets reveal specific disease mechanisms in IBD. Furthermore, we identify IBD-related pathways, ontologies, and cellular processes that are comparable between mouse models and patient cohorts. Our findings provide a valuable resource for selecting the most appropriate experimental paradigm to model unique features of IBD pathogenesis, allowing analysis at the tissue, cellular, and subcellular levels.

Introduction

Inflammatory bowel disease (IBD) represents a group of debilitating and chronic gastrointestinal disorders that include Crohn's disease (CD) and ulcerative colitis (UC). Despite significant progress in our understanding of IBD, the precise etiology remains elusive. Global incidence rates have been rising steadily for several decades. With an average age of onset of 30 years, the disease adds significant socioeconomic burden [1]. Moreover, incidence rates of pediatric IBD are also increasing, urging further research into better therapeutic measures [2]. It is hypothesized that IBD results from multiple triggers, including a) an exaggerated and dysregulated immune response, b) barrier dysfunction and microbial dysbiosis, c) genetic susceptibility, and d) other unknown environmental causes [3, 4, 5, 6]. Recent genome-wide association studies have identified more than 200 IBD-associated genomic loci, some of which directly confer increased susceptibility to IBD [7, 8, 9]. Moreover, recent evidence points to a potential role of autoantigens, elevated levels of which in some instances predate IBD onset [10, 11, 12]. However, direct causal relationships have only been established for a few genetic variants, and the contributions of other factors in triggering IBD has suffered from the “chicken or the egg” dilemma. This, together with the lack of shared molecular, prognostic, and etiological factors, has led investigators to hypothesize that IBD is not a single, defined disease, but a spectrum of diverse underlying disease mechanisms that result in overlapping clinical presentation [13].

To unravel the mechanisms that underlie the complex nature of IBD, various experimental mouse models of intestinal inflammation have been established. Modeling a complex disease in preclinical models is challenging, yet provides researchers with the opportunity to simulate disparate aspects of the disease in a controlled context [14, 15, 16, 17]. Experimental mouse models of intestinal inflammation are broadly based on three main approaches: (I) disruption of the intestinal epithelial barrier, (II) activation of an aberrant immune response, and (III) alteration of microbial homeostasis through infection. Each of these approaches replicates specific components of disease etiology, thereby providing invaluable tools for exploring the heterogeneity of IBD. However, none of these models fully reflect the complexity of human disease. This is exemplified by the distinct portrayal of particular pathobiological characteristics of disease in each experimental model. For instance, even the macroscopic endpoints, such as diarrhea, weight loss, fecal blood loss, site of inflammation, spleen-to-body weight ratio, colonoscopy score, and fibrosis vary between experimental models. Deciphering disease biology and the development of novel therapeutic strategies therefore relies on the choice of models to investigate the endophenotypes of concern.

To leverage preclinical modelling of intestinal inflammation and expose the paradigms for conserved IBD relevant molecular features we generated transcriptomic data from thirteen commonly used murine models of experimental intestinal inflammation. By employing both correlative statistics and curation-based measures of comparison across the regulomes of each mouse model we uncovered shared and unique co-regulated modules related to ciliary processes, inflammatory response, carboxylic acid metabolism, mitochondriopathy, and synaptic processes. Comparative transcriptomics between mouse models and multiple IBD datasets identified age-specific signatures, suggesting the existence of conserved regulatory mechanisms that are differentially affected in pediatric versus adult patients. Our findings offer a framework for comparing and selecting an appropriate modeling paradigm that best aligns with conserved disease processes, paving the way for discovering novel translatable therapeutic targets.

Results

Modeling the effects of diverse upstream triggers of inducing experimental intestinal inflammation

Experimental approaches of intestinal inflammation for modeling IBD in mice are broadly classified into three categories: barrier damage, immune modulation, and infection. To explore disease mechanisms, we recreated mouse models representing these categories on a C57BL/6J strain. We established cohorts for thirteen models (Fig. 1A): AcDSS and cDSS colitis [18], *Casp8*$^{\Delta IEC}$Ile and *Casp8*$^{\Delta IEC}$Col ileitis and colitis [19] for barrier damage; TC [20], OxC [18], AcTNBS and cTNBS [18], *Tnf*$^{\Delta ARE}$Ile and *Tnf*$^{\Delta ARE}$Col [21] for immune modulation; Everm [22], Hhepa [23], and Crode [24] for infection. For an overview of the models and abbreviations, please refer to Supplemental Table ST1. Models were screened using colonoscopy, in vivo imaging, and fecal bacterial load measurement (Supplemental Fig. S1A, B). Post-euthanasia histology and transgene expression provided tissue-level insights (Supplemental Fig. S1A, B).

We created a publicly accessible transcriptomic database and webtool denoted SEPIA (http://trr241.hosting.rrze.uni-erlangen.de/SEPIA/), from colon and ileum tissues of each model, enabling standardized comparative analyses (Supplemental Fig. S1C). The location of the tissue, rather than the inflammatory trigger, accounted for most variation, prompting separate analyses of colonic versus ileal models (Supplemental Fig. S1D). Next, we used scRNA-Seq datasets [25] and deconvolution algorithms [26] to generate signature matrices and extract cellular compositional features from bulk transcriptomes. In inflamed mice, higher immune cell proportions were evident, with distinct cellular compositions across models, including macrophage and neutrophil infiltration, colonocyte subtype loss, and stromal population impacts concurrent to histological and endoscopic findings of epithelial erosion, hypertrophy, and enhanced granularity (Fig. 1B, Supplementary Fig.S1A-B, and Table ST2).

New and unexpected insights included the detection of squamous-like epithelia in cDSS and *Casp8*$^{\Delta IEC}$Col, muscle-like cell reduction in AcTNBS, Hhepa, and Crode, and elevated B-cell signatures in AcDSS, cDSS, and AcTNBS models (Fig. 1B, Supplemental Table ST2). A poor representation of certain cell types such as enteric neuroglia, prompted us to take a curated approach for validating the deconvolution results. For this, we analyzed marker gene expression representing specific cell populations (Fig. 1C-D, Supplemental Fig S2A-B). These analyses confirmed the deconvolution data such as loss of secretory goblet cells in AcDSS, AcTNBS, Hhepa and Crode models, and revealed new findings such as an elevation in the transit amplifying (TA) cell markers in the Hhepa, Crode, and cTNBS colitis models, correlating with epithelial hypertrophy (Fig. 1C-D, Supplemental Fig S1B). Furthermore, we

detected that enteric neuronal marker levels contrasted strongly between infectious versus the AcDSS and OxC models (Fig. 1C-D, Supplemental Fig S1B). An interesting inverse correlation was detected between B-cell versus TA-cell markers, and a similar inverse trend emerged between enteric neuron versus TA-cell markers (Fig. 1D). Corresponding trends at the protein level confirmed our transcriptomic findings (Fig. 1E).

In small intestinal models, we observed model-specific immune cell marker changes (Supplemental Fig. S2A, B). Strikingly, the *Tnf*$^{\Delta ARE}$Ile model exhibited a strong inverse trend between T- and B-cell markers and was the only model with strong B-cell marker expression. Similar to colonic models, a negative correlation between the B-cell and TA-cell markers was also observed. Goblet cell markers were repressed in all small intestinal models, matching histological findings (Supplemental Fig S1A, S2A, B). In addition, the *Casp8*$^{\Delta IEC}$Ile model, showed only mild macrophage marker elevation, T-cell suppression, and no B-cell marker alteration despite histological inflammation (Supplemental Fig S1A, S2A, B).

Our analyses demonstrated that transcriptomic tissue alterations largely align with histological and preclinical metrics, revealing significant differences in cellular architecture within and between model categories, informing model choice when investigating specific cellular contexts.

Transcriptomic correlations among models reveal overlapping intra- and inter-category regulation

Having identified how different inflammatory onsets influence gut cellular composition, we conducted transcriptome-wide analyses of similarities and differences. Transcriptome-transcriptome correlation among the colonic and ileal models revealed the proportion of intra- and inter-category regulatory overlap (Fig. 2 and Supplemental Fig S2C). Models with the greatest transcriptomic impact were AcDSS (barrier damage), cTNBS (immune modulation), and Hhepa (infectious) (Fig. 2, diagonal pie charts). Within barrier damage models, AcDSS and cDSS showed moderate correlation of determination (R^2 = 0.65). Interestingly, cDSS and *Casp8*$^{\Delta IEC}$Col had the highest correlation within the barrier damage category (R^2 = 0.67), suggesting similar regulatory changes induced by chronic DSS administration and persistent IEC necroptosis in *Casp8*$^{\Delta IEC}$Col (Fig. 2). Some of the immune-driven colitis models also reached high correlation (OxC versus cTNBS, R^2 = 0.79) (Fig. 2). However, the highest intra-category correlation was within the infectious colitis category (R^2 = 0.86), despite the differing requirements for IL-10 receptor inhibition in the Hhepa versus the Crode models. (Fig. 2).

Besides intra-category correlations, several high-level inter-category correlations were also identified ($R^2 > 0.7$), notably between the barrier model *Casp8*$^{\Delta IEC}$Col and the infection model Crode ($R^2 = 0.8$) and the immune model TC ($R^2 = 0.7$) (Fig. 2). The infection model Crode showed some of the highest inter-category correlations against the immune models OxC and TC ($R^2 = 0.76$ and 0.71, respectively) (Fig. 2).

Interestingly, despite the size of the regulated transcriptome in the small intestinal models being fairly large (Supplemental Fig. S2C, diagonal pie charts), correlations remained low. The highest correlation was between *Casp8*$^{\Delta IEC}$Ile and *Tnf*$^{\Delta ARE}$Ile ($R^2 = 0.52$) (Supplemental Fig. S2C), indicating distinct and poorly concordant transcriptomic responses to inflammation triggers in the small intestine.

These analyses reveal contrasts in different model categories of gut inflammation. While anticipated similarities within each category were observed, significant cross-category similarities suggest common regulatory overlaps regardless of the nature of inflammatory onset.

Common and unique transcriptomic signatures emerge from diverse upstream triggers and simulate distinct molecular pathways of IBD

To directly reveal transcriptomic similarities and differences in each category, we evaluated gene ontology (GO) similarities among commonly co-expressed genes in colonic (Fig. 3A, upregulated; 3E, downregulated) and small intestinal models (Fig. 3B, upregulated; 3F, downregulated). We performed ontology-based semantic similarity analysis to determine functional attributes and relatedness between these genes [27, 28, 29]. Commonly upregulated genes in colonic models confirmed expected inflammatory responses, forming three clusters: innate immune response, chemotaxis, and response to bacteria (Fig. 3A, C). In contrast, small intestinal models showed upregulation in rRNA metabolic process, RNA processing, and telomere organization (Fig. 3B, D).

The 279 commonly downregulated genes in colonic models (Fig. 3E) didn't enrich against known ontology terms, except for Wnt signaling (Fig. 3G). In stark contrast, the 389 commonly downregulated genes in the ileal models (Fig. 3F), were enriched for alpha-defensins indicating suppressed antimicrobial responses (Fig. 3H). Detailed ontology terms, adjusted p-values and kappa coefficients are provided in Supplemental Table ST3. In addition to the shared genes and processes, we also analyzed the unique genes specific to

each model identifying model-specific ontologies (Supplemental Fig. S3A-D; Supplemental Table ST3).

Next, we investigated the relationship between models and hallmark pathways associated with IBD using gene set variation analysis (GSVA) coupled with limma-based estimation of differences [30, 31, 32]. Nine distinct hallmark pathways were selected from the MySigDB collection, relevant to inflammation and epithelial homeostasis [33, 34]. The enrichment scores for specific pathways were highly similar for models within each category. In colonic barrier damage models, pathways like inflammatory response, IFNγ response, and TNF signaling were robustly activated. Whereas in the colonic immune modulation category, DNA repair and TNF signaling pathways were activated (Fig. 3I). Crode and Hhepa infection models showed concordance in upregulated pathways, except for Wnt-β-catenin signaling, which was reduced in Crode but unaffected in Hhepa. Notch and TGBβ signaling were upregulated in Hhepa but not in Crode (Fig. 3I). The AcTNBS and OxC models showed reduced Notch and Wnt-β-catenin signaling (Fig. 3I). It was striking that the TC model of immune modulation showed pathway enrichment similar to colonic infection models, with significant enrichment in most pathways, including TNF signaling (Fig. 3I).

Among the small intestinal models, the TNF signaling reached significance only in the *Tnf*$^{\Delta ARE}$Ile model, with a trend in the Everm, but not the *Casp8*$^{\Delta IEC}$Ile model (Fig. 3I). The IFNγ response pathway was significantly enriched among all three small intestinal models (Fig. 3I). These analyses provide a reference for selecting appropriate models for investigating specific pathways in IBD and developing interventions.

Identification of co-expression-based gene modules across mouse models of intestinal inflammation

To circumvent the biases that arise from pathway annotation and database selection, we employed weighted correlation network analysis (WGCNA) [35] to identify gene co-expression modules across our mouse model datasets. Our analysis yielded 16 and 11, colonic and ileal modules of co-expression, respectively (Fig. 4A, B; Supplemental Table ST4-5). The analysis yielded expected results, demonstrating the co-expression of specific modules in both colonic as well as ileal models. For instance, modules ME11 and ME4, respectively, were identified in both models and consisted of GO biological processes related to “inflammatory responses” (Fig. 4C, D). Similarly, the modules ME3 (colon) and ME1 (ileum) exhibited a

preponderance of negative Δ eigenvalues and were composed of ontologies related to "carboxylic acid processes" (Fig. 4C, D; Supplementary table ST6).

Interestingly, the 437 colonic (Fig. 3A) and 468 ileal (Fig. 3B) commonly upregulated genes identified earlier, fell into the ME11 (colon) and ME4 (ileum) modules, respectively. These modules consisted of immune regulatory ontologies with genes such as *Nod2*, *Nfkbia, Il1b*, *Il18r1*, *Lcp2*, *Itgb2*, among others (Fig. 4C-E; Supplemental Table ST4-6).

Moreover, our analysis revealed that ME3 (colon) and ME1 (ileum) consisted of ontologies belonging to the carboxylic acid metabolic process, with notable regulatory genes associated with immune suppression. For example, these modules included the gene *Nr5a2*, which encodes for LRH-1, a protein involved in the generation of immunomodulatory corticosteroids from cholesterol. They also included the genes *Cyp2w1*, *Cyp39a1*, and *Cyp2c55,* which control cholesterol catabolism and supply for the synthesis of corticosteroids [36]. It is well established that both host-derived and commensal carboxylic acid species are known to regulate immune responses. Interestingly, the fact that the same modules also contain genes that regulate the supply of monocarboxylic acids, including, for example, *Acsm3*, *Acss2, Ppargc1a*, and *Acsl2,* which are involved in the control of metabolic immunomodulation (Fig. 4C,E; Supplemental Table ST4,6) [37, 38, 39]. It was unexpected that the same genes representing these pathways were also found to be repressed in human IBD, as described later in the following sections (Fig. 5C, D). This suggests that a generalized suppression of a co-regulated network of endogenous immunosuppression and modulation across inflamed tissues may occur, regardless of the initial trigger of inflammation or its location.

In addition to the aforementioned common modules, tissue- and model-specific co-regulatory gene ontologies with opposing Δ eigenvalues were also identified. For example, ontologies pertaining to 'neuronal processes' were observed in ME4 (colon), which exhibited a specific and inverse correlation in the AcDSS model (Fig. 4A). Similarly, in the ileal models, ME2 consisted of unique gene ontologies for 'RNA processing' with positive Δ eigenvalues for the genes *Wdr36*, *Dck1*, *Utp4*, and *Wdr75*, which control both pre- and mature ribosomal RNA processing. These genes also play a role in the context of p53 stress-response [40, 41, 42, 43] (Fig. 4F; Supplemental Table ST5-6).

Comparative transcriptomics identify conserved and divergent co-expression ontologies between preclinical models and IBD patients

Comparative transcriptomics enables comparing patient and preclinical models, facilitating the identification of conserved disease mechanisms and preclinical testing [44, 45]. Using public and our own IBD-patient transcriptomes (Supplemental table ST7), we compared them with our mouse model data to analyze commonly regulated ontologies relevant to IBD [46, 47]. This revealed a number of interesting correlations with moderate to high R^2 for specific IBD-relevant KEGG pathways, stratified by disease location: colonic (Fig. 5A) and ileal (Fig. 5B). The core pathways driving inflammation, such as chemokine signaling, cytokine receptor signaling, the JAK/STAT pathway, and TNF signaling, showed higher R^2 values across various model model-human cohort pairs (Fig. 5 A, B, and Supplemental Table ST8).

Specific KEGG pathways showed interesting correlations between model–cohort pairs. For instance, in specific IBD cohorts, glutathione metabolism, cell adhesion, and antigen processing and presentation exhibited a high goodness of fit and met significance thresholds against infectious colitis models (Fig. 5A). However, among the small intestinal models, several orthologs did not reach the expected expression-enrichment thresholds (Fig. 5B, ‘x’ marks). A distinctive correlation pattern was observed for the pathway of gap junctions in the AcDSS model versus the PROTECT and IBDome patient cohorts. The Th1-Th2 differentiation pathway showed significant goodness of fit for the Hhepa and AcDSS models versus most IBD cohorts (Fig. 5A, Supplemental table ST7).

The Th17 differentiation pathway is an important player in the pathogenesis of IBD. This inflammatory pathway of Th17 cell differentiation showed a somewhat unexpected trend, with significant to high R^2 values. Among the colonic models, this pathway was significantly correlated in most of the model-cohort pairs, except the *Casp8*$^{\Delta IEC}$Col and TC models (Fig. 5A). Similarly, in the ileal models, Th17 cell differentiation was significant in most model-cohort pairs (5/6) where ortholog expression-enrichment thresholds were met (Fig. 5B).

The IBD cohorts in our study are broadly divided into pediatric (PROTECT and RISK) and adult (PSC-UC, WASH-U, and IBDome). We identified specific correlations unique to pediatric cohorts, such as glycolysis in the *Tnf*$^{\Delta ARE}$Ile and Everm models; TGFβ signaling in the TC and *Tnf*$^{\Delta ARE}$Ile models; JAK/STAT signaling in the TC model; and Th17 cell differentiation in the cDSS model (Fig. 5 A, B).

Interestingly, several of the genes from colonic modules ME11 and ME3 were also concordantly regulated in the PROTECT (UC) patient transcriptomes with similar magnitudes and directions (Fig. 5C). Most genes commonly regulated among ileal models (Fig. 3E, F) were also regulated in the RISK (CD) patient cohort (Fig. 5D). This regulatory overlap between the mouse models and CD patients exhibited a conserved direction and magnitude.

Furthermore, these commonly regulated genes from the mouse ileal models (Fig. 3E, F) were represented on the ileal co-expression modules, ME1 and ME2 (Fig. 4D, F). However, the specific ontology associated with RNA processing was not affected in the human RISK cohort (Fig. 5D, lack of highlighted datapoints). Thus, comparative transcriptomics between IBD patient samples and preclinical mouse models revealed conserved disease mechanisms that replicate in at least one of the mouse models in a given category. These included key pathways like chemokine signaling, JAK/STAT, and Th17 cell differentiation. On the other hand, regulatory behavior of specific pathways correlates better between a given model category and patient tissues. For example, glycolysis across the immune modulation and the infectious categories and glutathione metabolism and antigen presentation across the infectious categories. Moreover, there is a need for better barrier damage induced ileitis mouse models that are more representative of human CD-like ileitis. These findings highlight both shared and unique regulation of key IBD relevant pathways in model-cohort pairs.

Conserved and divergent ontologies follow unique patterns of expression in IBD-associated cell clusters

Recent single-cell RNA sequencing (scRNA-Seq) data has revealed new cellular states and subsets emerging during gut inflammation and exhibiting non-canonical expression profiles. For example, Li *et al*. identified a new specialized IEC subtype with the capacity to regulate mucosal immunity and drive human CD [48]. To ascertain whether specific inflammation-associated co-expression modules that we identified (Fig. 4), can be observed in non-canonical cell subsets that emerge in human IBD, we analyzed our data against scRNA-Seq studies from IBD patients.

Indeed, analysis of UC patient- scRNA-Seq datasets [49] revealed the emergence of cross-functional states of TA, stem, and immature enterocytes highly expressing genes in the colonic module ME11 enriched for the GO “regulation of immune response”. This GO is canonically ascribed to immune cell lineages (Fig. 5E). In addition, M-cells, which are critical for host defense, also showed an upregulation of this GO class in inflamed UC tissues (Fig. 5E).

Another novel cellular state, specific to inflamed patients was a plasma cell subcluster. This subcluster displayed an elevated expression of genes enriched for the GO term “angiogenesis”, as identified in the colonic module ME5 (Fig. 5F). Additionally, an elevated level of expression of the GO term “angiogenesis” was observed on the $CD69^{+}$ mast cells cluster specifically in inflamed UC tissues (Fig. 5F). Surprisingly, a recently discovered novel

stromal cluster of inflammatory fibroblasts and subclusters of endothelial and post-capillary venule endothelial cells exhibited an elevated expression of genes in colonic module ME3 enriched for the GO “carboxylic acid metabolic process” (Fig. 5G). We also performed an extended module-by-cell type analysis at the gene level for each colonic cell lineage (epithelial, immune, and stromal) and stratified by disease status. It was observed that there was a notable degree of cell type specificity in the expression of genes from various modules such as ME5 (goblet), ME8 (cycling TA), and ME14 (M cells) modules, in addition to some other differences in the average expression of modules based on the disease status. For example, expression of genes from ME14 was highly enriched in inflamed UC M-cells (Supplemental Fig. 4A; Supplementary Table ST9).

Among the ileitis mouse models, we identified “RNA processing” as an enriched GO class from the module ME2 (Fig. 4F), which did not align with any of the IBD cohort data. Moreover, the magnitude of expression for this GO class was comparable across clusters from the scRNA-Seq data of CD ileal samples [50] (Supplemental Fig. S5A, B). However, we detected specific modules derived from ileitis mouse models that exhibited cellular-level alterations when compared to the ileal CD scRNA-Seq data (Supplemental Fig. S5B; Supplemental Table ST9).

Collectively our analyses highlight how conserved expression of specific ontologies may result from the emergence of specific cellular states or subtypes during inflammation in the gut. Since the modules analyzed here were drawn from the commonly regulated transcripts across mouse models, it is apparent that different modalities of inflammation onset yield similar changes in cellular states in IBD.

Impact of inflammatory processes on cellular substructures are conserved in specific IBD cohorts and select mouse models.

Aberrant transcript expression and physical organization of IEC substructures, such as microvilli, has been reported to affect patients with CD suggestive of epithelial dysfunction [51]. Interestingly, analysis of our preclinical model data had also identified a common co-expression module, ME0, containing several ontologies related to “cilium” (Fig. 4C, D). Indeed, WGCNA modules with negative Δ eigengene scores for ontologies related to ‘cilium’ were enriched in the adult colonic PSC-UC (ME1) cohort [52] (Fig. 6A-C). A module consisting of ontologies related to “cilium” was also detected in the RISK ileum CD cohort, but it did not show any significant enrichment (Fig. 6A-C). We reconfirmed the original findings from the

WashU ileum CD cohort where ciliary dysfunction was first reported, validating the analysis approach (Fig. 6A-C). Confirming this, scanning electron micrographs of the epithelia from patients in our study cohort revealed microvilli damage and their disrupted organization, further supporting the notion that this feature is not only restricted to CD, but can also contribute to epithelial dysfunction in UC (Fig. 6D-E). However, it is noteworthy that none of the pediatric IBD cohorts, which included RISK and PROTECT, had modules containing enrichments for the term 'cilium' or related gene ontologies indicative of a strong age dependency in this subcellular feature. Moreover, cilium and related ontologies were significantly depleted in specific mouse models including the colitis models Crode, cTNBS, and AcTNBS and the ileitis models Everm and *Tnf*$^{\Delta ARE}$ (Fig. 6F-G). Our data indicate that damage to epithelial microvilli is common in CD as well as UC and affects adult more than pediatric IBD patients and this feature can be modelled in specific preclinical mouse models of intestinal inflammation.

Another subcellular organelle that is impacted in IBD is mitochondria. A recent study demonstrated how mitochondriopathy contributes to disease severity and treatment response in the UC PROTECT cohort [53, 54]. We replicated these findings in our in-house cohort where the expression levels of multiple mitochondrially encoded genes as well as the protein levels of COX4 were appreciably reduced in the UC patient samples (Fig. 6H-I). Interestingly, when we extrapolated these findings to our mouse model data, the mitochondrial changes were most strikingly modeled in the infectious models, Everm and Hhepa (Fig. 6J). It is noteworthy that in contrast to the infectious mouse models, the cTNBS colitis model showed an inverse correlation with respect to the expression of mitochondrially encoded genes versus that in UC patients.

Overall, our data highlight the conserved and divergent regulation of genes and pathways in preclinical intestinal inflammation models and clinical IBD cohorts. Utilizing these as a tool for pre-screening while targeting IBD disease processes will enable guided discovery efforts.

Discussion

Research efforts in preclinical modeling of disease pathways to understand the underlying mechanisms of ectopic and chronic inflammation in patients are hindered by the poor translatability of results. A common challenge is selecting an appropriate model to study the pathway and mechanism in question. Given the diversity of models and their modes of inflammation induction, this choice becomes even more complex. Our data summarize how

comparative transcriptomic analyses can elucidate shared and model-specific pathway regulation, as well as identify translatable components between preclinical models and clinical disease.

Our findings indicate that preclinical mouse models share commonly regulated determinants essential for the inflammatory processes observed in the gut, correlated with the anatomical site of inflammation. These fundamental regulatory overlaps, emerging from disparate triggers of preclinical intestinal inflammation, appear to be well-conserved in clinical IBD. Conversely, the identification of unique transcriptomic profiles in each mouse model will enable the preselection of appropriate models for investigating specific molecular mechanisms and pathways.

In conducting these analyses, we also generated comprehensive comparisons of the alterations occurring in the models at the tissue level tracking cellular changes using deconvolution approaches. The deconvolution approach is more robust than other methods because it does not rely on preselected cell type marker genes. Instead, it takes into account global transcriptomic differences between cell types derived from known changes occurring in single-cell transcriptomes. The advantage of this approach is that as new cell states and subtypes are identified by the scRNA-Seq studies, their detection on bulk data through deconvolution becomes feasible. Interestingly, Liu *et al*. recently showed that during colonic inflammation, the squamous epithelial cells from the transition zone migrate towards the proximal colon [25]. The presence of intermediate squamous-like cells was identified in the *Casp8*$^{\Delta IEC}$Col, cDSS, and AcTNBS, models and appeared to negatively associate with neutrophil signatures. It is tempting to speculate whether these intermediate squamous-like cells are partially resistant to injury and erosion and hence may reduce neutrophil infiltration [25, 55]. Another interesting inverse association was observed between TA-cells and B-cells. This finding is consistent with the recent observation that B-cell expansion can impede epithelial regeneration [56]. However, further studies are necessary to elucidate the precise interplay between epithelial regeneration and B-cells in the context of IBD.

One limitation of deconvolution approaches to construct cellular signatures from bulk tissue data is their reliance on well-annotated scRNA-Seq datasets to derive signature matrices. In this respect, we constrained our analysis using the publicly available scRNA-Seq dataset from DSS-induced colitis model. Therefore, the inclusion of additional scRNA-Seq data from various intestinal inflammation models would enhance and refine future deconvolution attempts with bulk transcriptomic data.

By employing the analytical findings derived from the mouse model data and comparing them with IBD patient datasets, we were able to recapitulate the regulatory behavior of conserved pathways in IBD. Two well-established and publicly available IBD transcriptomic datasets derived from the pediatric PROTECT and RISK cohorts were subjected to this analysis [53, 57]. In addition, we conducted a comparative analysis of biopsy-derived transcriptomes from three distinct patient cohorts; our IBDome cohort of UC and CD patients, a PSC-UC cohort [52], and an additional cohort of adult CD patients (WashU cohort) [51]. Interestingly, downregulation of the immunomodulatory pathways regulated by *NR5A2, CYP2W1, ACSS2, ACSM3, and PPARGC1A* was conserved among the mouse model and the PROTECT and to some extent, the RISK cohorts. In addition, the conserved upregulation of several core inflammatory genes, including *NOD2, IL18R1, IL1B,* and *ITGB2*, between colonic mouse models and the pediatric IBD cohorts mentioned above. Moreover, the recent report of a mitochondriopathy-like phenotype was also observed in the pediatric PROTECT and RISK cohorts. This phenotype was most evident in the infectious colitis models but not in cTNBS [53, 57]. This may indicate that the chronic activation of proinflammatory pathways in IBD, at the expense of immunometabolic pathways, may impact treatment response. Indeed, specific processes were associated with the adult IBD cohorts, which did not recapitulate in the pediatric cohorts. For example, the identification of the downregulation of the GO processes ciliary organization and elevated microvilli damage was conserved among the adult IBD cohorts of PSC-UC and WashU and the mouse models [51].

By superimposing signatures from specific co-expression modules on scRNA-Seq data from immune cell lineages of IBD patients, we also identified a novel cellular state of plasma cells with lymphangiogenic functions in the inflamed tissues. These findings are consistent with the recent identification of a new B-cell subset with angiogenic properties by van de Veen *et al*. [58]. Yet another novel, inflammation associated cellular subtype the elusive inflammatory fibroblasts that emerges during active flares in IBD patients, is functionally ill-defined. Some data indicate a proinflammatory role, whereas other data suggest a potential regenerative function [44]. The elevated average expression of the carboxylic acid metabolic process in inflammatory fibroblasts was intriguing. The GO class carboxylic acid metabolic process, which controls the supply of a plethora of endogenous metabolites implicated in immune cell regulation [59, 60], demonstrated an overall repression in our mouse model data as well as in IBD transcriptomes. Specific stromal clusters from inflamed tissues of UC patients, including inflammatory fibroblasts, post-capillary venule endothelia, and glia, showed an overall increase in the expression signature for the carboxylic acid metabolic process, which is involved in producing lipid intermediates and endogenous corticosteroids. A recent hypothesis suggests that during inflammation and cancer, stromal cells may act as a source

of gut-specific immunosuppressive mediators, such as PGE2 [15, 61, 62, 63]. However, this may also reflect high metabolic demands on these cells, driving the production of large quantities of extracellular matrix molecules or metabolites derived from carboxy acids. The functional significance of inflammatory fibroblasts in controlling or promoting inflammation remains unclear and requires further study. Whether such immunomodulatory signaling is predominantly paracrine or whether niche-specific cells such as post-capillary venule endothelia and glia may also contribute to immunomodulation need further investigation.

In summary, we present a resource for the rapid profiling of global transcriptomic alterations in preclinical intestinal inflammation models and provide comparative references against human cohorts. The observations made in our study have important implications for choosing mouse models to investigate disease-driving mechanisms in IBD. Finally, our analyses provide proof of concept for the composite, overlapping, and conserved core program of inflammation in IBD emerging from apparently diverse triggers. Our identification of unique model-specific processes highlights potential context-specific regulatory features. Further investigation and replication of these features may reveal unique molecularly-defined IBD subclusters. This will enable a guided choice of mouse models, to investigate the intended disease mechanism in light of its regulatory conservation in human disease.

Materials and methods

Mice

All mice on a C57BL6/J background were purchased from Charles River (Charles River GmbH) and Janvier Labs (Janvier Labs GmbH) and were housed under specific pathogen-free (SPF) conditions.

Acute Dextran Sulfate Sodium (DSS)-induced colitis

Five wild-type mice with a C57BL6/J background were subjected to 2% DSS with molecular weight ranging from 36,000 – 50,000 Da (MP Biomedicals) dissolved in drinking water for eight days. Evaluation of symptoms of inflammation and colitis parameters was performed as indicated in previous studies [64]. The mice were sacrificed on day nine to obtain inflamed colonic tissue for RNA sequencing.

Chronic DSS-induced colitis

C57BL6/J mice were subjected to three one-week (7 day) cycles of repeated 2% DSS of molecular weight ranging from 36,000 – 50,000 Da (MP Biomedicals) interspersed by 2 two-week stages of recovery. Mice were euthanized 10 days post-the last DSS cycle, and their colon was collected for RNA sequencing of the tissue.

Oxazolone-induced colitis

The induction of inflammation in the oxazolone model requires a preliminary sensitization step, whereby 3% (wt/vol) solution of 4-Ethoxymethylene-2-phenyl-2-oxazolin-5-one (oxazolone) (Sigma-Aldrich) in a 1:4 dilution with olive oil / acetone mixture to the shaved skin of the animal, as previously described [18]. Seven days following the sensitization procedure, intrarectal enema administration of 150µl of 0.5% oxazolone in 50% ethanol was conducted. Mice were sacrificed 24 hours later and the colonic tissues were dissected and snap frozen in liquid nitrogen using cryovials for RNA sequencing and fixed in 4% Histofix (Karl Roth GmbH) for histology.

Acute TNBS-induced colitis

Colitis was induced by the hapten TNBS (2,4,6-trinitrobenzenesulfonic acid, Merck) as previously described [18]. Mice were sensitized by epicutaneous application of 1% TNBS at a dilution of 1:4 in a mixture of oil and acetone (100 µl) on day 0, followed by intrarectal administration of 2% TNBS in 45% ethanol (100 µl) on day 6, followed by monitoring of body

weight and mini-endoscopy. The mice were euthanized and colonic tissues were harvested and snap frozen in cryovials for RNA Sequencing and fixed in 4% Histofix (Karl Roth GmbH) for histology.

Chronic TNBS-induced colitis

For the induction of chronic colitis by TNBS, mice were sensitized as above [18] with 1% of TNBS in a mixture of oil and acetone (1:4). Subsequently, mice were administered with 150 µl of 0.5% TNBS via rectal enema on a 14-day schedule over a three-cycle period. Weight was monitored at various time points. The development of colitis was monitored over time by means of mini-endoscopy. At the end of the experiment, mice were euthanized and colonic tissues were harvested, snap frozen and fixed as above.

T-cell transfer colitis

For this model, a special control group of immune-deficient *Rag1*$^{-/-}$ mice, coming from a C57BL6/J background is necessary. For the inflammation model, naïve CD4^{+} CD25^{-} T-cells were isolated from the spleen of wild-type C57BL6/J mice using MACS-based isolation (CD4^{+} T Cell Isolation Kit, mouse, Miltenyi Biotec) with additional CD25 negative selection (CD25 Antibody, anti-mouse, Miltenyi Biotec). The obtained cells were controlled for their viability and purity via flow cytometry. One million cells were injected intraperitoneally in *Rag1*$^{-/-}$ mice on a C57BL6/J background. The development of colitis was monitored over time via assessment of weight loss and mini-coloscopy. Colonic tissue for RNA sequencing was extracted from the sacrificed mice 3 weeks after the inoculation.

***Caspase8*$^{\Delta IEC}$ ileitis and colitis**

Casp8$^{\Delta IEC}$ mice were generated by crossing mice carrying a loxP-flanked caspase-8 allele (*Casp8*fl) mice to Villin-Cre mice, which were described earlier [19]. Cre-mediated recombination was genotyped by polymerase chain reaction on tail DNA The development of ileitis and colitis has been previously characterized [19]. The onset of ileitis was observed to occur approximately at week 12, with the mice being sacrificed at approximately week 14. Mouse disease burden was assessed regularly according to the FELASA guidelines.

***Tnf*$^{\Delta ARE}$ ileitis**

The *Tnf*$^{\Delta ARE}$ mouse model is a widely used genetic model for studying Crohn's disease. *Tnf*$^{\Delta ARE}$ mice are characterized by the deletion of the AU-rich element (ARE) in the tumor necrosis factor-alpha (TNF) gene, resulting in the constitutive overexpression of TNF. This

leads to the development of chronic tissue inflammation, particularly affecting the terminal ileum. The *Tnf*$^{\Delta ARE}$ mouse model closely resembles human Crohn's disease in terms of its histological features and the shared pathogenetic role of TNF. Mice heterozygous for the *Tnf*$^{\Delta ARE}$ modification start to develop histological abnormalities at the age of 6 weeks [21]. For this study, 12-week-old heterozygous mice with fully established signs of ileal inflammation were used. The mouse line was a kind gift by Fabio Cominelli (Case Western Reserve University, USA).

***Eimeria vermiformis* infection**

Oocysts from *E. vermiformis* (provided by Marc Veldhoen (Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Portugal) were obtained and stored in 2.5% potassium bicromate, according to previously published protocols [65]. The oocysts were washed three times with deionized water, centrifuged at 1,800 g for 8 minutes, and floated at 1,100 g for 10 minutes. A final sterilization was performed using sodium hypochlorite. After three more washing cycles with deionized water, they were counted with the help of a Fuchs-Rosenthal chamber. One thousand of the cleaned oocysts diluted in water were used for infection of the mice by oral gavage. The mice were sacrificed nine days after the infection, and the ileal tissue was collected for RNA sequencing.

***Helicobacter hepaticus* infection**

H. hepaticus (DSMZ no.: 22909) was revived in glycerol stocks and cultured on blood agar plates under microaerobic conditions at 37°C. Microscopic examination confirmed the bacterial quality and purity before the cultures were transferred to tryptone soya broth supplemented with 10% fetal calf serum, 10 µg/mL vancomycin, 5 µg/mL trimethoprim, and 2.5 IU/mL polymyxin B. Cultivation proceeded under microaerophilic conditions at 37°C, with shaking at 180 rpm. Mice were orally inoculated with 1 × 108 colony-forming units (c.f.u.) of *H. hepaticus* on two consecutive days using a 22 G curved blunted needle. Concurrently, mice received intraperitoneal injections (i.p.) of 1 mg of anti-IL-10R antibody (clone 1B1.2, BioXcell) on day 0 and at weekly intervals thereafter [23, 66, 67]. On day 14 post-inoculation, mice were euthanized, and colon samples were harvested for subsequent analyses. Sections of ~0.3 cm from colon were preserved in RNA Later (Qiagen) for gene expression studies. Histological evaluation of colitis was carried out in accordance with the methodology previously described [68].

***Citrobacter rodentium* infection**

C. rodentium is a Gram-negative bacterium that mimics human enteropathogenic E. coli (EPEC) infections. For this model we used the ICC169 strain of *C. rodentium*, as referenced in previous studies [69]. The organism was cultivated in a sterilized LB medium at a temperature of 37°C under constant aeration and shaking conditions. The culture was supplemented with erythromycin. For the inoculation of the bacteria, mice were subjected to eight hours of fasting, after which ~10^9 colony-forming units (CFU) of *C. rodentium,* diluted in sterile phosphate-buffered saline (PBS), were introduced into the animal via oral gavage. Eight days later, the mice were sacrificed and their colons were collected for RNA sequencing.

Experimental protocols were approved by the Institutional Animal Care and Use Committee of the Regierung von Unterfranken and the Landesamt für Gesundheit und Soziales – Berlin, and in accordance with the UK Scientific Procedures Act of 1986.

Human cohort data

Publicly available IBD patient cohorts that were used in this study include the treatment naïve cohorts GSE109142 (PROTECT), GSE117993 (RISK-UC) [53], GSE57945 (RISK - CD) [57], cohort of adult IBD with concurrent PSC E-MTAB 7915 (PSC-UC) [52], and the adult Crohn's disease patient cohort E-MTAB 5783 (WashU) (50, 51). For the in-house cohort (IBDome), samples were obtained from non-IBD and IBD patients after written informed consent was obtained from the participants at the Gastroenterology Department, of Charite – Universitätsmedizin Berlin as well as the Universitätsklinikum Erlangen. Biopsies were taken during endoscopy or removed from resected tissue and stored in 10 mL RNA protect Tissue Reagent (Qiagen) overnight at 4°C. For long-term storage, the reagent was removed and biopsies kept at -80°C. For RNA isolation, the RNeasy Kit (Qiagen) was used according to the manufacturer´s instructions employing the TissueLyser LT (Qiagen) and 5 mm balls. Concentration and purity of the isolated RNA was determined by the ratios of absorbance (A260:280 and A260:230) using NanoDrop 1000 (Thermo Fisher Scientific). The RNA Integrity Number (RIN) was determined using the Agilent RNA ScreenTape System (Agilent Technologies) according to manufacturer's instructions employing the Agilent 2200 TapeStation. The RNA Clean & Concentrator™-25 (Zymo Research) was used for RNA samples with a RIN < 1.8 according to the manufacturer´s instructions. RNA samples were shipped on dry ice to Quantitative Biology Center (QBiC) of Eberhard Karls Universität Tübingen for sequencing. All data from the patients were pseudonymized.

Histological and immunohistological analyses

For the hematoxylin and eosin (H&E) stainings, we prepared sections of 5-10 µm from tissues fixed in formalin and embedded in paraffin blocks, which were then subjected to standard H&E staining protocol. The immunostainings were performed on paraffin sections of the tissues, which were cut, deparaffinized, hydrated, and treated with a Tris-EDTA-based antigen retrieval solution. To prevent non-specific binding, a commercial blocking reagent (Immunoblock 1X, Carl Roth GmbH) was utilized. The samples were then incubated overnight at 4°C in the dark with the primary antibody for each immunostaining (table below). We washed the samples in TBS and then we incubated them with either the secondary antibody or streptavidin conjugates (DyLight, Invitrogen) for 1 h at 4°C. Samples then were counterstained with Hoechst 33342 (Thermo Fisher Scientific) for the cellular nuclei and cover-slipped in fluorescence mounting medium. A Leica TCS SP5 confocal microscope (Leica Microsystems) was employed for the acquisition of the immunofluorescence images using the necessary settings.

Antibody	**Type**	**Dilution**	**Species**	**Code**	**Company**
Primary antibody					
Anti-F4/80	Unconjugated	1:200	Rabbit	7007S	Cell Signaling
Anti-CD3	Unconjugated	1:100	Rat	100238	Biolegend
Anti-KI67	Alex Fluo555	1:100	Rat	14-5698-82	abcam
Anti-MUC2	Unconjugated	1:200	Rabbit	NBP1-31231	Novus B
Anti-S100B	Unconjugated	Ready to use	Rabbit	Dako	Biolegend
Anti-TUBB3	Unconjugated	1:100	Mouse	801212	Biolegend
Anti-COX4		1:2400	Rabbit	4844	Cell Signaling
Anti-TOMM20		1:200	Rabbit	HPA011562	Sigma
Secondary antibody					
Anti-Rabbit	Biotinylated	1:500	Goat	111-065-144	Jackson
Anti-Rat	Biotinylated	1:500	Goat	554014	BD Biosciences

RNA extraction

The RNA obtained from the collected tissue samples was extracted using multiple extraction kits, according to the manufacturer's instructions, using the peqGold total tissue TNA kit

(peqlab GmbH) and the microspin total RNA kit (VWR International GmbH). The obtained RNA was controlled for contamination and degradation with the NanoPhotometer® spectrophotometer (IMPLEN, CA, USA), the Nanodrop (thermofischer), the Qbit (Thermo), and the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

Library Preparation and mRNA Sequencing

We aliquoted 1 µg of RNA per sample to prepare the libraries for sequencing. We used the NEBNext® UltraTM RNA Library Prep Kit for Illumina® (NEB, USA) according to the manufacturer's instructions, with indexing added to the molecules from each sample. From the total RNA of each sample, mRNA was purified using poly-T oligo-attached magnetic beads. The molecules were fragmented with divalent cations at high temperatures while diluted in NEBNext First Strand Synthesis Reaction Buffer (5X). We used a random hexamer primer and M-MuLV reverse transcriptase (RNase H-) to synthesize the first strand of cDNA, and DNA polymerase I was followed to synthesize the second strand. Excess RNA was cleared using RNase H. The overhang of the molecules was blunted through the action of the exonuclease/polymerase. The 3' ends of the DNA fragments were adenylated and a NEBNext adaptor with a hairpin loop structure was ligated to them to prepare for hybridization. The AMPure XP system (Beckman Coulter, Beverly, USA) was used to purify the fragments of 150-200 bp in length. To prepare the fragments for PCR, we used 3 µl of USER Enzyme (NEB, USA) at 37°C for 15 minutes and 5 minutes at 95°C. We performed the PCR with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer. The products of the PCT were purified again with the AMPure XP system and the resulting library was controlled for quality using the Agilent Bioanalyzer 2100 system.

We used a cBot Cluster Generation System using PE Cluster Kit cBot-HS (Illumina) for the clustering of the indexed samples according to the manufacturer's instructions. The prepared libraries were sequenced using an Illumina platform, generating paired-end reads. Initial trimming of the adapter and poly-N sequences, as well as demultiplexing of the reads was done by the Illumina software included in the platform. Further processing of the raw data (raw reads) was done using in-house scripts and publicly available software. Further analysis was performed in R software for statistical programming (4.4.0). Quality control of the fastq files was performed using FastQC (v0.12.1), calculating Q20, Q30 and GC content. We downloaded the current mouse reference genome (GRCm39) and its annotation from NCBI, UCSC and Ensembl. The genome was indexed and used for mapping the paired-end reads using STAR mapping software (2.7.10b) followed by sequence alignments using picard (3.1.1) and samtools (1.20) and the analysis of counts per gene using featureCounts (v2.0.6).

Analysis of ontology based semantic-similarity networks

To uncover the functional attributes and the relatedness among these shared genes, we performed an analysis for ontology-based semantic similarity on these genes (24-26). This method yields a network of sub-clusters where the nodes are ontologies and the edges are the kappa similarity coefficients. These analyses were performed as described by Sanchez *et al*. in their original article [29]. Clustering was performed on all terms reaching significance and with a kappa similarity of above 0.3. Cytoscape (3.10.2) was used for the visualization of the similarity network.

Deconvolution of gene expression

Processing of publicly available scRNA-Seq datasets was performed using scanpy (1.10.1) in python (3.10.12), using the provided data and annotation. The DWLS package (0.1.0) was used for the cell proportion deconvolution analysis. GSVA (1.52.2) was employed for the hallmark pathway analysis, with the genes extracted from GSEA database. Differential analyses of resulting values were performed using limma (3.60.2).

Analysis of differential gene expression

Differential expression analysis was performed using the DESeq2 R package (1.44.0), filtering out genes with less than 15 counts overall. Over representation analyses (ORA) using the Gene Ontology and KEGG enrichment databases was performed using the ClusterProfiler R package (4.12.0).

Weighted gene correlation network analysis

We used the WGCNA package (1.72-5) [35] on the raw counts to obtain co-expression modules across samples, segregated by tissue. Default parameters and a soft threshold was chosen as maximum for every analyses. A post-clustering comparison of these module eigenvalues using limma-based linear modeling was performed as shown previously [70], yielding the differential changes per module across treatment groups, which we represented in our study as Δ eigenvalues. An ORA was performed using the database Gene Ontology to observe processes represented in each module using the ClusterProfiler R package (4.12.0).

Analysis of publicly available scRNA-Seq data

Orthogonal analysis of human IBD scRNA-Seq datasets (UC accession No. SCP259 and CD accession No SCP1423) to identify cell clusters expressing the transposed mouse module data was performed using scanpy (1.10.1) in python (3.10.12).

Generation of figure plots

The commercial software Graphpad Prism 9 and in-house python (3.11.0rc1) and R (4.4.0) scripts were used to generate the graphical displays. R package ComplexHeatmap (2.20.0) was used for the generation of heatmaps, Ridge plots were generated using the ggridges package (0.5.6). Other plots in R were generated using the ggplot2 package (3.5.1). Python package pyvenn was used for the generation of the venn diagrams. Matplotlib (3.9.0) and Seaborn (0.13.2) were used for the figure plots generated using python (3.11.0rc1). Cytoscape (3.10.2) was used for network data visualization.

Ancillary packages and dependencies used in data processing, transformation and management:

Package	**Version**
Python packages	
Pandas	2.2.2
numpy	1.26.4
scipy	1.13.1
snakemake	7.32.4
flask	3.0.3
mysql.connector	8.4.0
reportlab	4.2.0
adjustText	1.1.1
R packages	
dendextend	1.17.1
cluster	2.1.6
gplots	3.1.3.1
circlize	0.4.16
gsubfn	0.7
GOstats	2.70.0
enrichplot	1.24.0
pathview	1.44.0
simplifyEnrichment	1.14.0

magrittr	2.0.3

Figure legends:

Figure 1. Changes in tissue cell composition across mouse models of IBD. A) Scheme depicting the mouse model data bank for preclinical gut inflammation, highlighting the distribution of the broad categories: infection, barrier damage, and immune modulation **B**) Deconvolved cell type signatures across colonic models. Inflamed samples are shown next to controls. The columns Ctrl_OxC and Ctrl_Crode are shared reference controls for the two respective models. **C**) Bubble plot showing fold change and adjusted p values of selected cell type markers across the colonic mouse models. **D**) Ridge plots of the fold change of selected cell type marker genes across the colonic mouse models. **E**) Immunofluorescence staining for markers for macrophages (green, first column), goblet cells (Green, second column), transit amplifying cells (red, second column), enteric neurons (green, third column), and enteric glia (red, third column) across the indicated models from each category.

Figure 2. Correlations between colonic mouse models. Pie charts on the diagonal represent the number of genes that are significantly up- (orange) and down- (violet) regulated in each model and the legend with the empty bubbles indicates the number of genes. Scatter plots in the bottom left half show the linear regression between two given models, comparing the log2-fold change of common genes. The color intensity of individual dots correlates directly with the slope of the regression line. Dot plots on the top right half represent the R^2 value of the respective linear regression. Triangular outlines demarcate the comparisons within a given model category, namely barrier damage, immune modulation, or infection.

Figure 3. Functional analyses of common features between models. A) Venn diagram showing the common upregulated genes between the indicated colitis models and **B)** those between ileitis models. **C-D**) Semantic similarity networks from the enriched GO terms observed in the upregulated gene set from the common datasets defined in A and B respectively. **E**) Venn diagrams showing the common downregulated genes between the indicated colitis models and **F)** those between the ileitis models. **G-H**) Term similarity network of the enriched GO terms observed in the downregulated gene set from the common datasets defined in E and F respectively**. I**) Variation in the relative expression of deconvolved hallmark pathways per mouse model showing variation profiles across the model dataset.

Figure 4. Weighted Gene Co-expression Network Analysis (WGCNA) across the mouse model dataset. A-B) Variation observed in the modules obtained from the WGCNA applied to colitis and ileitis samples from the mouse model dataset. Changes labeled with indicate post-LIMMA p values of < 0.05. **C-D**) Top three Gene Ontologies enriched in selected modules from the WGCNA of colitis (C) and ileitis (D) respectively. **E-F**) Heatmap of the enriched genes from the selected WGCNA modules from the colitis and ileitis mouse models respectively.

Figure 5. Translation of mouse model features to human IBD. A-B) Correlation sets for selected ontologies between (A) multiple mouse colitis models versus human UC cohort transcriptomes and (B) mouse ileitis models versus human CD cohort transcriptomes. Color intensity represents the R^2 value. Framed cells denote a significant p-value in the regression test ($p < 0.05$). Cells marked with an [X] had less than 10 genes in common and the regression was not performed. **C**) Volcano plots of the indicated UC patient cohort where (left) orthologs from the common upregulated genes from colitis mouse models and (right) downregulated genes are shown. All gene commonly regulated among mouse colitis models are highlighted and those belonging to the WGCNA mouse colitis modules ME11 and ME3 are labelled in red **D**) Volcano plots of the indicated CD patient cohort where (left) orthologs from the common upregulated gene set from ileitis mouse models and (right) downregulated genes are shown. All genes commonly regulated among mouse models are highlighted and those belonging to the WGCNA mouse ileitis modules ME2 and ME1 are labeled in red. **E-G**) t-SNEs of single cell experiment from epithelial, immune, and stromal compartments of the human gut by Smillie et. al. [49] t-SNE plots are segregated by patient health and color represents the average gene expression of selected pathways from the WGCNA modules. Violin plots show the changes of the average expression in each of the annotated clusters.

Figure 6. Microvilli dysfunction and mitochondriopathy observed in IBD translate to selected mouse colitis and ileitis models. A) Variation observed in the mouse colitis and ileitis WGCNA modules applied to the indicated human cohorts. Significant changes labelled with * indicating p values of < 0.05 in the differential analyses. **B)** Top 7 ontologies enriched in selected modules from the analysis in A). C) Volcano plots of the indicated human cohorts with genes from the selected WGCNA modules highlighted in orange and those from the specific enriched ontologies from B) in purple. **D)** Electron micrographs of the intestinal epithelial cells from healthy individuals. **E)** Electron micrographs of epithelial cells from UC patients **F-G)** Bubble plot showing the variation in the relative expression of selected ontologies from the mouse colitis F) and ileitis G) models. **H)** Volcano plots from the indicated UC and CD patient transcriptomes with mitochondrially expressed genes. **I)** Immunofluorescence staining for mitochondrial marker COX4 in healthy individuals and UC

patient colonic samples. **J)** Bubble plot showing the log2 fold change in the expression of mitochondrially encoded genes across all mouse models of colitis and ileitis.

Supplemental Figure legends

Supplementary Figure 1. Clinical features of mouse models. A) Representative H&E histology images of the ileitis mouse models and healthy controls along with the indicated pre- and post-euthanasia endpoints for disease establishment. **B)** Representative images from pre-euthanasia screening of inflammation (top row: colonoscopy and IVIS Spectrum *in vivo* imaging), and post-euthanasia histological assessments (bottom row) from the colitis mouse models from each model category. **C**) Schematic illustration of the processing and analysis of the mouse-mouse and the mouse-human datasets **D)** Principal component analyses (PCA) of the samples in the model databank (grey ellipse = colonic samples, yellow ellipse = ileal samples), few individual models have been highlighted.

Supplementary Figure 2. Tissue composition across mouse models. A) Bubble plot showing the expression of selected cell type markers across the ileitis mouse models. **B**) Ridge plots of the selected cell type markers across the ileitis mouse models. **C**) Correlations between ileal mouse models. Pie charts on the diagonal represent the number of genes that are significantly up- (orange) and down- (violet) regulated in each model and the legend with the empty bubbles indicates the number of genes. Scatter plots in the bottom right half show the linear regression between two given models, comparing the log2-fold change of common genes. The color of the dots is directly proportional to the slope of the regression line. Dot plots on the top left half represent the R^2 value of the linear regression.

Supplementary Figure 3: Exclusive processes observed in mouse models. A-B) Exclusively up- (A) and down- (B) regulated genes from the colitis mouse models. The top three enriched Gene Ontologies are shown for each model. **C-D**) Exclusively up- (C) and down- (D) regulated genes from ileitis mouse models. The top three enriched Gene Ontologies are shown for each model.

Supplementary Figure 4. Average expression of WGCNA modules across different single-cell clusters from UC samples. **A-C**) Violin plots of average expression of the genes from each WGCNA module identified from the colitis models among the epithelial (A),

immune (B), and stromal (C) compartments, segregated by annotated cell type and patient health.

Supplementary Figure 5. Translation of WGCNA modules and features from mouse ileitis models onto single-cell RNAseq from CD patients. A) t-SNE of single cell RNAseq PREDICT experiment (left) and average expression of genes from the ontology 'RNA processing' included in mouse ileitis module ME2, segregated by inflammation status. **B**) Violin plots of average expression of the genes identified from each mouse ileitis WGCNA module segregated patient health and annotated by cell type.

References

1. Ng SC, *et al*. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet* **390**, 2769-2778 (2017).

2. Kuenzig ME, *et al*. Twenty-first Century Trends in the Global Epidemiology of Pediatric-Onset Inflammatory Bowel Disease: Systematic Review. *Gastroenterology* **162**, 1147-1159 e1144 (2022).

3. Ananthakrishnan AN, *et al*. Lifestyle, behaviour, and environmental modification for the management of patients with inflammatory bowel diseases: an International Organization for Study of Inflammatory Bowel Diseases consensus. *Lancet Gastroenterol Hepatol* **7**, 666-678 (2022).

4. Neurath MF. Host-microbiota interactions in inflammatory bowel disease. *Nat Rev Gastroenterol Hepatol* **17**, 76-77 (2020).

5. Neurath MF. Targeting cytokines in inflammatory bowel disease. *Sci Transl Med* **14**, eabq4473 (2022).

6. Friedrich M, Pohin M, Powrie F. Cytokine Networks in the Pathophysiology of Inflammatory Bowel Disease. *Immunity* **50**, 992-1006 (2019).

7. Sazonovs A, *et al*. Large-scale sequencing identifies multiple genes and rare variants associated with Crohn's disease susceptibility. *Nat Genet* **54**, 1275-1283 (2022).

8. Liu JZ, *et al*. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* **47**, 979-986 (2015).

9. de Lange KM, *et al.* Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* **49**, 256-261 (2017).

10. Kuwada T, *et al.* Identification of an Anti-Integrin alphavbeta6 Autoantibody in Patients With Ulcerative Colitis. *Gastroenterology* **160**, 2383-2394 e2321 (2021).

11. Lundgren BA, *et al.* Identification of complement C3 as an autoantigen in inflammatory bowel disease. *Eur J Gastroenterol Hepatol* **22**, 429-436 (2010).

12. Livanos AE, *et al.* Anti-Integrin alphavbeta6 Autoantibodies Are a Novel Biomarker That Antedate Ulcerative Colitis. *Gastroenterology* **164**, 619-629 (2023).

13. Uhlig HH, Powrie F. Translating Immunology into Therapeutic Concepts for Inflammatory Bowel Disease. *Annu Rev Immunol* **36**, 755-781 (2018).

14. Uhlig HH, *et al.* Differential activity of IL-12 and IL-23 in mucosal and systemic innate immune pathology. *Immunity* **25**, 309-318 (2006).

15. Patankar JV, *et al.* E-type prostanoid receptor 4 drives resolution of intestinal inflammation by blocking epithelial necroptosis. *Nat Cell Biol* **23**, 796-807 (2021).

16. Neurath MF, Fuss I, Kelsall BL, Stuber E, Strober W. Antibodies to interleukin 12 abrogate established experimental colitis in mice. *J Exp Med* **182**, 1281-1290 (1995).

17. Duckworth CA. Understanding gene-environment interactions in a mouse model of Crohn's disease. *Dis Model Mech* **4**, 7-8 (2011).

18. Wirtz S, *et al.* Chemically induced mouse models of acute and chronic intestinal inflammation. *Nat Protoc* **12**, 1295-1309 (2017).

19. Gunther C, *et al.* Caspase-8 regulates TNF-alpha-induced epithelial necroptosis and terminal ileitis. *Nature* **477**, 335-339 (2011).

20. Ostanin DV, *et al.* T cell-induced inflammation of the small and large intestine in immunodeficient mice. *Am J Physiol Gastrointest Liver Physiol* **290**, G109-119 (2006).

21. Kontoyiannis D, Pasparakis M, Pizarro TT, Cominelli F, Kollias G. Impaired on/off regulation of TNF biosynthesis in mice lacking TNF AU-rich elements: implications for joint and gut-associated immunopathologies. *Immunity* **10**, 387-398 (1999).

22. Roberts SJ, *et al.* T-cell alpha beta + and gamma delta + deficient mice display abnormal but distinct phenotypes toward a natural, widespread infection of the intestinal epithelium. *Proc Natl Acad Sci U S A* **93**, 11774-11779 (1996).

23. Fox JG, Ge Z, Whary MT, Erdman SE, Horwitz BH. Helicobacter hepaticus infection in mice: models for understanding lower bowel inflammation and cancer. *Mucosal Immunol* **4**, 22-30 (2011).

24. Borenshtein D, McBee ME, Schauer DB. Utility of the Citrobacter rodentium infection model in laboratory mice. *Curr Opin Gastroenterol* **24**, 32-37 (2008).

25. Liu CY, *et al.* Transitional Anal Cells Mediate Colonic Re-epithelialization in Colitis. *Gastroenterology* **162**, 1975-1989 (2022).

26. Sturm G, *et al.* Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics* **35**, i436-i445 (2019).

27. Gu Z, Hubschmann D. simplifyEnrichment: A Bioconductor Package for Clustering and Visualizing Functional Enrichment Results. *Genomics Proteomics Bioinformatics* **21**, 190-202 (2023).

28. Zhou Y, *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523 (2019).

29. Sanchez D, Batet M. Semantic similarity estimation in the biomedical domain: an ontology-based information-theoretic perspective. *J Biomed Inform* **44**, 749-759 (2011).

30. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417-425 (2015).

31. Ritchie ME, *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).

32. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7 (2013).

33. Castanza AS, Recla JM, Eby D, Thorvaldsdottir H, Bult CJ, Mesirov JP. Extending support for mouse data in the Molecular Signatures Database (MSigDB). *Nat Methods* **20**, 1619-1620 (2023).

34. Subramanian A, *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).

35. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).

36. Fernandez-Marcos PJ, Auwerx J, Schoonjans K. Emerging actions of the nuclear receptor LRH-1 in the gut. *Biochim Biophys Acta* **1812**, 947-955 (2011).

37. Yi Q, *et al.* Ascl2 facilitates IL-10 production in Th17 cells to restrain their pathogenicity in inflammatory bowel disease. *Biochem Biophys Res Commun* **510**, 435-441 (2019).

38. Miller KD, *et al.* Acetate acts as a metabolic immunomodulator by bolstering T-cell effector function and potentiating antitumor immunity in breast cancer. *Nat Cancer* **4**, 1491-1507 (2023).

39. Yao Y, Cai X, Fei W, Ye Y, Zhao M, Zheng C. The role of short-chain fatty acids in immunity, inflammation and metabolism. *Crit Rev Food Sci Nutr* **62**, 1-12 (2022).

40. Peluso AA, *et al.* Quantitative Phosphoproteomics of the Angiotensin AT(2)-Receptor Signaling Network Identifies HDAC1 (Histone-Deacetylase-1) and p53 as Mediators of Antiproliferation and Apoptosis. *Hypertension* **79**, 2530-2541 (2022).

41. Skarie JM, Link BA. The primary open-angle glaucoma gene WDR36 functions in ribosomal RNA processing and interacts with the p53 stress-response pathway. *Hum Mol Genet* **17**, 2474-2485 (2008).

42. Wilkins BJ, Lorent K, Matthews RP, Pack M. p53-mediated biliary defects caused by knockdown of cirh1a, the zebrafish homolog of the gene responsible for North American Indian Childhood Cirrhosis. *PLoS One* **8**, e77670 (2013).

43. Moudry P, Chroma K, Bursac S, Volarevic S, Bartek J. RNA-interference screen for p53 regulators unveils a role of WDR75 in ribosome biogenesis. *Cell Death Differ* **29**, 687-696 (2022).

44. Kinchen J, *et al.* Structural Remodeling of the Human Colonic Mesenchyme in Inflammatory Bowel Disease. *Cell* **175**, 372-386 e317 (2018).

45. Gonzalez Acera M, *et al.* Comparative Transcriptomics of IBD Patients Indicates Induction of Type 2 Immunity Irrespective of the Disease Ideotype. *Front Med (Lausanne)* **8**, 664045 (2021).

46. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* **27**, 29-34 (1999).

47. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30 (2000).

48. Li J, *et al.* A Specialized Epithelial Cell Type Regulating Mucosal Immunity and Driving Human Crohn's Disease. *bioRxiv*, (2023).

49. Smillie CS, *et al.* Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* **178**, 714-730 e722 (2019).

50. Zheng HB, *et al.* Concerted changes in the pediatric single-cell intestinal ecosystem before and after anti-TNF blockade. *medRxiv*, 2021.2009.2017.21263540 (2023).

51. VanDussen KL, *et al.* Abnormal Small Intestinal Epithelial Microvilli in Patients With Crohn's Disease. *Gastroenterology* **155**, 815-828 (2018).

52. Quraishi MN, *et al.* A Pilot Integrative Analysis of Colonic Gene Expression, Gut Microbiota, and Immune Infiltration in Primary Sclerosing Cholangitis-Inflammatory Bowel Disease: Association of Disease With Bile Acid Pathways. *J Crohns Colitis* **14**, 935-947 (2020).

53. Haberman Y, *et al.* Ulcerative colitis mucosal transcriptomes reveal mitochondriopathy and personalized mechanisms underlying disease severity and treatment response. *Nat Commun* **10**, 38 (2019).

54. Pelia R, *et al.* Profiling non-coding RNA levels with clinical classifiers in pediatric Crohn's disease. *BMC Med Genomics* **14**, 194 (2021).

55. Leppkes M, *et al.* Neutrophils prevent rectal bleeding in ulcerative colitis by peptidyl-arginine deiminase-4-dependent immunothrombosis. *Gut* **71**, 2414-2429 (2022).

56. Frede A, *et al.* B cell expansion hinders the stroma-epithelium regenerative cross talk during mucosal healing. *Immunity* **55**, 2336-2351 e2312 (2022).

57. Haberman Y, *et al.* Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. *J Clin Invest* **124**, 3617-3633 (2014).

58. van de Veen W, *et al.* A novel proangiogenic B cell subset is increased in cancer and chronic inflammation. *Sci Adv* **6**, eaaz3559 (2020).

59. Artyomov MN, Van den Bossche J. Immunometabolism in the Single-Cell Era. *Cell Metab* **32**, 710-725 (2020).

60. Barnhoorn MC, Hakuno SK, Bruckner RS, Rogler G, Hawinkels L, Scharl M. Stromal Cells in the Pathogenesis of Inflammatory Bowel Disease. *J Crohns Colitis* **14**, 995-1009 (2020).

61. Brown SL, *et al.* Myd88-dependent positioning of Ptgs2-expressing stromal cells maintains colonic epithelial proliferation during injury. *J Clin Invest* **117**, 258-269 (2007).

62. Miyoshi H, *et al.* Prostaglandin E2 promotes intestinal repair through an adaptive cellular response of the epithelium. *EMBO J* **36**, 5-24 (2017).

63. Roulis M, *et al.* Paracrine orchestration of intestinal tumorigenesis by a mesenchymal niche. *Nature* **580**, 524-529 (2020).

64. Becker C, Fantini MC, Neurath MF. High resolution colonoscopy in live mice. *Nat Protoc* **1**, 2900-2904 (2006).

65. Figueiredo-Campos P, Ferreira C, Blankenhaus B, Veldhoen M. Eimeria vermiformis Infection Model of Murine Small Intestine. *Bio Protoc* **8**, (2018).

66. Kullberg MC, *et al.* Helicobacter hepaticus-induced colitis in interleukin-10-deficient mice: cytokine requirements for the induction and maintenance of intestinal inflammation. *Infect Immun* **69**, 4232-4241 (2001).

67. West NR, *et al.* Oncostatin M drives intestinal inflammation and predicts response to tumor necrosis factor-neutralizing therapy in patients with inflammatory bowel disease. *Nat Med* **23**, 579-589 (2017).

68. Izcue A, *et al.* Interleukin-23 restrains regulatory T cell activity to drive T cell-dependent colitis. *Immunity* **28**, 559-570 (2008).

69. Riedel CU, Casey PG, Mulcahy H, O'Gara F, Gahan CG, Hill C. Construction of p16Slux, a novel vector for improved bioluminescent labeling of gram-negative bacteria. *Appl Environ Microbiol* **73**, 7092-7095 (2007).

70. Liu Y. CWGCNA: an R package to perform causal inference from the WGCNA framework. *NAR Genom Bioinform* **6**, lqae042 (2024).

Data availability

All of the raw sequencing data generated from this study has been uploaded to the European nucleotide archive via ArrayExpress (https://www.ebi.ac.uk/biostudies/arrayexpress) and is available for download using the following accession numbers (AcDSS, cDSS, and TC: E-MTAB-14306; Everm: E-MTAB-14297; Hhepa: E-MTAB–14316; OxC, Crode: E-MTAB–14312; *Casp8*$^{\Delta IEC}$Col and *Casp8*$^{\Delta IEC}$Ile: E-MTAB-14318; AcTNBS, cTNBS: E-MTAB-14329; *Tnf*$^{\Delta ARE}$Col and *Tnf*$^{\Delta ARE}$Ile: E-MTAB-14325). We have also created an interactive webtool that enables accessing and browsing of processed data that can be accessed using the following link http://trr241.hosting.rrze.uni-erlangen.de/SEPIA/ (user login available upon request). Publicly available IBD cohorts can be accessed at the gene expression omnibus https://www.ncbi.nlm.nih.gov/geo/ (accession IDs- PROTECT: GSE109142, RISK CD: GSE57945, and RISK UC: GSE117993) and ArrayExpress (accession IDs - PSC-UC: E-MTAB-7915 and WashU: E-MTAB-5783). Access to the IBDome dataset is available upon request. Publicly available mouse single cell RNA sequencing datasets used in this study can be accessed using the following accession number: GSE168033. Accession for human scRNA-Seq datasets used in this study are UC SCP259 and CD SCP1423.

Code availability

All code generated in this study is has been deposited on the following git hub repository and is accessible via the following link https://github.com/MiguelGonzalezAcera/SEPIA .

Acknowledgements

We acknowledge the support provided for animal husbandry and rearing for this work by Eva Wellein and Melanie Ziegler. Parts of this manuscript are under preparation for being submitted as part of the thesis leading to the academic title doctor of philosophy for MGA.

Funding support

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - TRR241 375876048 (A03, A05, A08, B05, Z03,), SFB1181 (C05), and individual grants with project numbers 418055832, 531560612, and 510624836. The project was further supported by CRU 5023 (project number 50474582), CRC 1449-B04 and Z02

(project number 431232613); CRC 1340-B06 (project number 372486779). A.N.H is supported by a Lichtenberg fellowship and “Corona Crisis and Beyond” grant by the Volkswagen Foundation, a BIH Clinician Scientist grant and INST 335/597-1, as well as with the ERC-StG “iMOTIONS” grant (101078069). Moreover, the project was supported by the DFG grant covering sequencing costs to B.S and C.B Projektnummer 418055832. The project was also supported by the Interdisciplinary Centre for Clinical Research (IZKF: A76, A93, J96, and ELAN P120).

Author Contributions

The study was conceived by MGA, JVP, and CB and planned BS, ANH, CB, MN, ZT, and the IBDome Consortium. Data was analyzed by MGA, JVP, CP, and GS. The figures were prepared by MB, MGA, and JVP. MGA and JVP wrote the manuscript. Help with manuscript editing was provided by MGA, JVP, ANH, BS, CB, ZT, RGB, KH, LE, MB, and AAK. Data was acquired by CP, GS, LE, RC, RGB, TL, LB, MD, RW, IS, KG, FM, and KK. Data interpretation was performed by MGA, JVP, HL, ANH, and CB. Help with securing funds JVP, CB, MN, BS, ANH, SW, BW, CG, KH, and the IBDome Consortium. Study Supervision JVP, KH, ZT, SW, BW, and CB.

Other contributing authors

TRR241 IBDome Consortium: Imke Atreya[1], Raja Atreya[1], Petra Bacher[2,3], Christoph Becker[1], Christian Bojarski[4], Nathalie Britzen-Laurent[1], Caroline Bosch-Voskens[1], Hyun-Dong Chang[5], Andreas Diefenbach[6], Claudia Günther[1], Ahmed N. Hegazy[4], Kai Hildner[1], Christoph S. N. Klose[6], Kristina Koop[1], Susanne Krug[4], Anja A. Kühl[4], Moritz Leppkes[1], Rocío López-Posadas[1], Leif S.-H. Ludwig[7], Clemens Neufert[1], Markus Neurath[1], Jay V. Patankar[1], Magdalena Prüß[3], Andreas Radbruch[5], Chiara Romagnani[3], Francesca Ronchi[6], Ashley Sanders[4,8], Alexander Scheffold[2], Jörg-Dieter Schulzke[4], Michael Schumann[4], Sebastian Schürmann[1], Britta Siegmund[4], Michael Stürzl[1], Zlatko Trajanoski[9], Antigoni Triantafyllopoulou[5,10], Maximilian Waldner[1], Carl Weidinger[4], Stefan Wirtz[1], Sebastian Zundler[1]

[1]Department of Medicine 1, Friedrich-Alexander University, Erlangen, Germany

[2]Institute of Clinical Molecular Biology, Christian-Albrecht University of Kiel, Kiel, Germany.

[3]Institute of Immunology, Christian-Albrecht University of Kiel and UKSH Schleswig-Holstein, Kiel, Germany.

[4]Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Gastroenterology, Infectious Diseases and Rheumatology, Berlin, Germany

[5]Deutsches Rheuma-Forschungszentrum, ein Institut der Leibniz-Gemeinschaft, Berlin, Germany

[6]Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Institute of Microbiology, Infectious Diseases and Immunology

[7]Berlin Institute für Gesundheitsforschung, Medizinische System Biologie, Charité – Universitätsmedizin Berlin

[8]Max Delbrück Center für Molekulare Medizin, Charité – Universitätsmedizin Berlin

[9]Biocenter, Institute of Bioinformatics, Medical University of Innsbruck, Innsbruck, Austria.

[10]Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Rheumatology and Clinical Immunology, Berlin, Germany

Figure 1

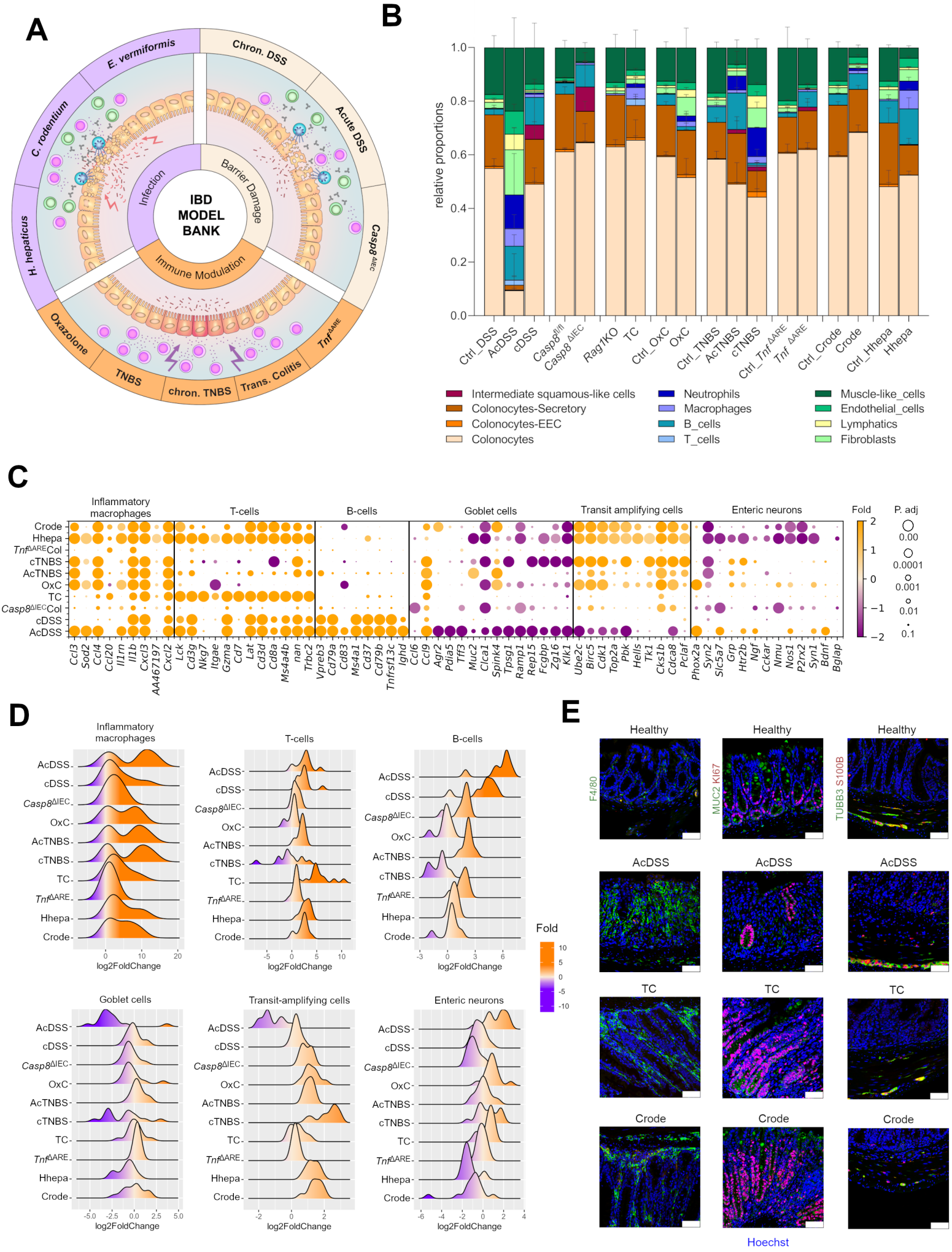

A
E. vermiformis
Chron. DSS
C. rodentium
Acute DSS
H. hepaticus
Casp8ΔIEC
Oxazolone
Tnf ΔARE
TNBS
chron. TNBS
Trans. Colitis
Infection
Barrier Damage
Immune Modulation
IBD MODEL BANK
B
relative proportions
Ctrl_DSS
AcDSS
cDSS
Casp8 fl/fl
Casp8 ΔIEC
Rag1KO
TC
Ctrl_OxC
OxC
Ctrl_TNBS
AcTNBS
cTNBS
Ctrl_Tnf ΔARE
Tnf ΔARE
Ctrl_Crode
Crode
Ctrl_Hhepa
Hhepa
Intermediate squamous-like cells
Colonocytes-Secretory
Colonocytes-EEC
Colonocytes
Neutrophils
Macrophages
B_cells
T_cells
Muscle-like_cells
Endothelial_cells
Lymphatics
Fibroblasts
C
Inflammatory macrophages
T-cells
B-cells
Goblet cells
Transit amplifying cells
Enteric neurons
Fold
P. adj
D
log2FoldChange
E
Healthy
AcDSS
TC
Crode
F4/80
MUC2 KI67
TUBB3 S100B
Hoechst

Figure 2

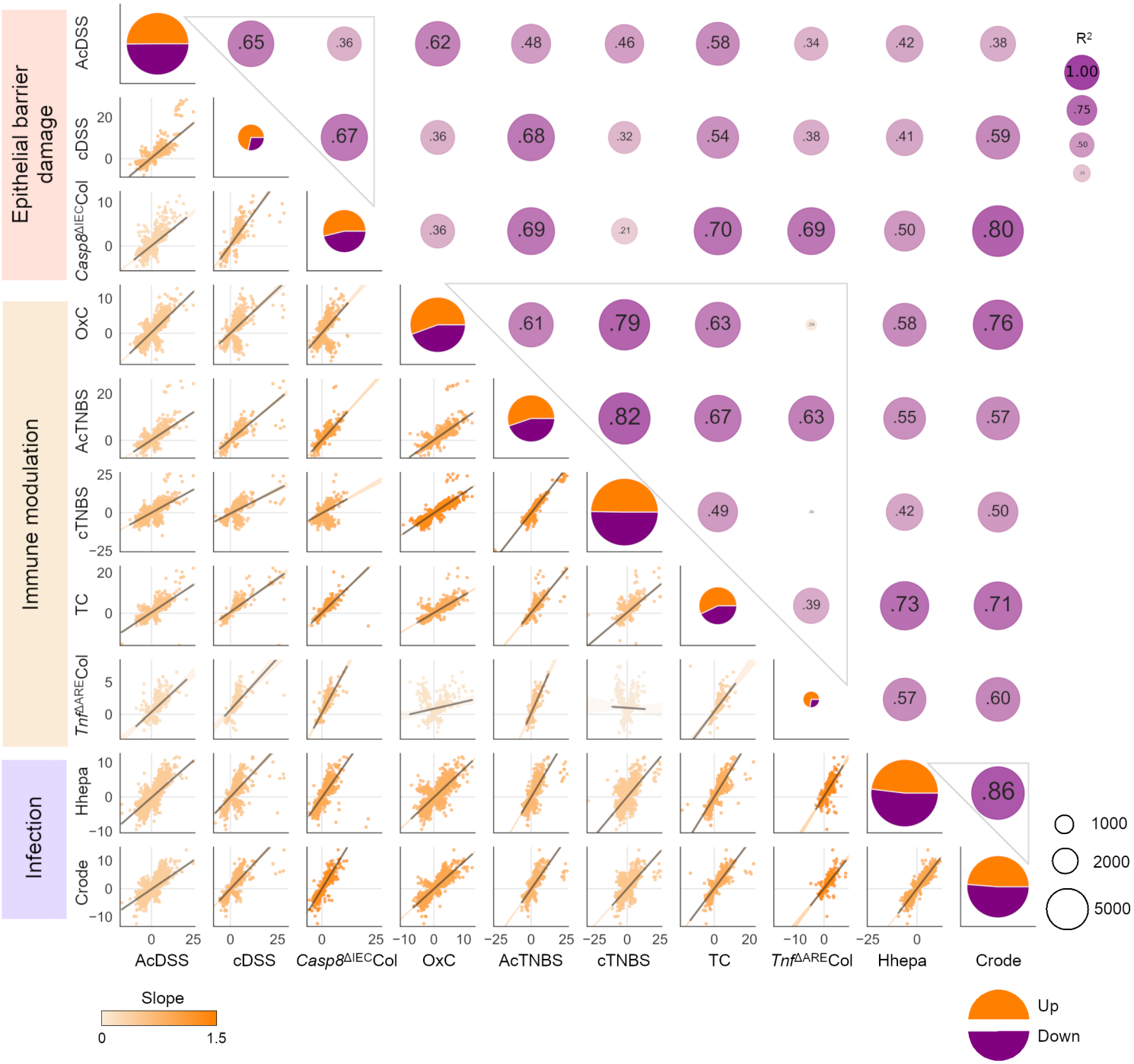

Epithelial barrier damage
Immune modulation
Infection
AcDSS
cDSS
Casp8ΔIECCol
OxC
AcTNBS
cTNBS
TC
TnfΔARECol
Hhepa
Crode
R2
1.00
.75
.50
.25
1000
2000
5000
Slope
0
1.5
Up
Down

Figure 3

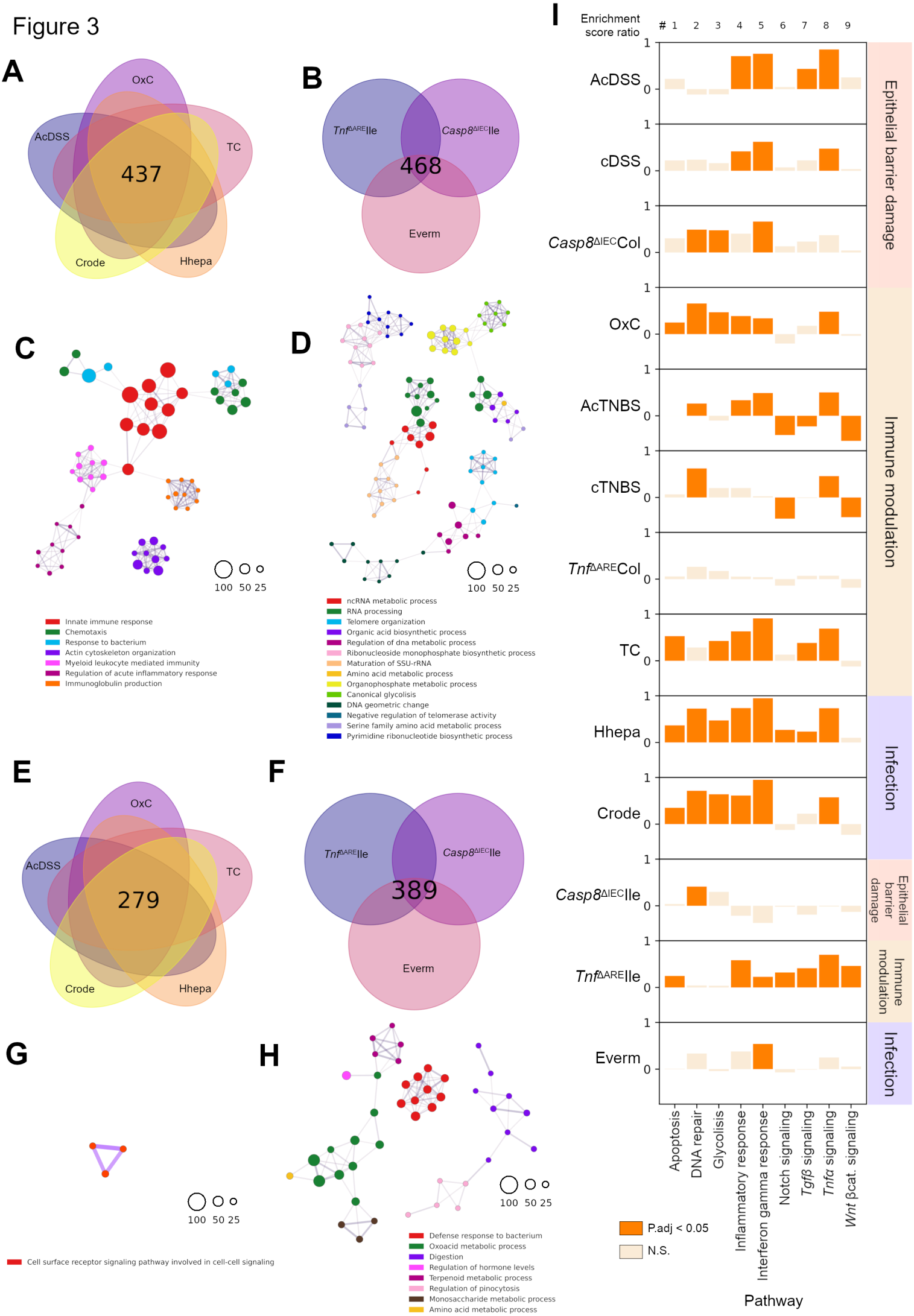

A
OxC
AcDSS
TC
437
Crode
Hhepa
B
TnfΔAREIle
Casp8ΔIECIle
468
Everm
C
100 50 25
Innate immune response
Chemotaxis
Response to bacterium
Actin cytoskeleton organization
Myeloid leukocyte mediated immunity
Regulation of acute inflammatory response
Immunoglobulin production
D
100 50 25
ncRNA metabolic process
RNA processing
Telomere organization
Organic acid biosynthetic process
Regulation of dna metabolic process
Ribonucleoside monophosphate biosynthetic process
Maturation of SSU-rRNA
Amino acid metabolic process
Organophosphate metabolic process
Canonical glycolisis
DNA geometric change
Negative regulation of telomerase activity
Serine family amino acid metabolic process
Pyrimidine ribonucleotide biosynthetic process
E
OxC
AcDSS
TC
279
Crode
Hhepa
F
TnfΔAREIle
Casp8ΔIECIle
389
Everm
G
100 50 25
Cell surface receptor signaling pathway involved in cell-cell signaling
H
100 50 25
Defense response to bacterium
Oxoacid metabolic process
Digestion
Regulation of hormone levels
Terpenoid metabolic process
Regulation of pinocytosis
Monosaccharide metabolic process
Amino acid metabolic process
I
Enrichment score ratio
1 2 3 4 5 6 7 8 9
AcDSS
cDSS
Casp8ΔIECCol
OxC
AcTNBS
cTNBS
TnfΔARECol
TC
Hhepa
Crode
Casp8ΔIECIle
TnfΔAREIle
Everm
Epithelial barrier damage
Immune modulation
Infection
Epithelial barrier damage
Immune modulation
Infection
Apoptosis
DNA repair
Glycolisis
Inflammatory response
Interferon gamma response
Notch signaling
Tgfβ signaling
Tnfα signaling
Wnt βcat. signaling
P.adj < 0.05
N.S.
Pathway

Figure 4

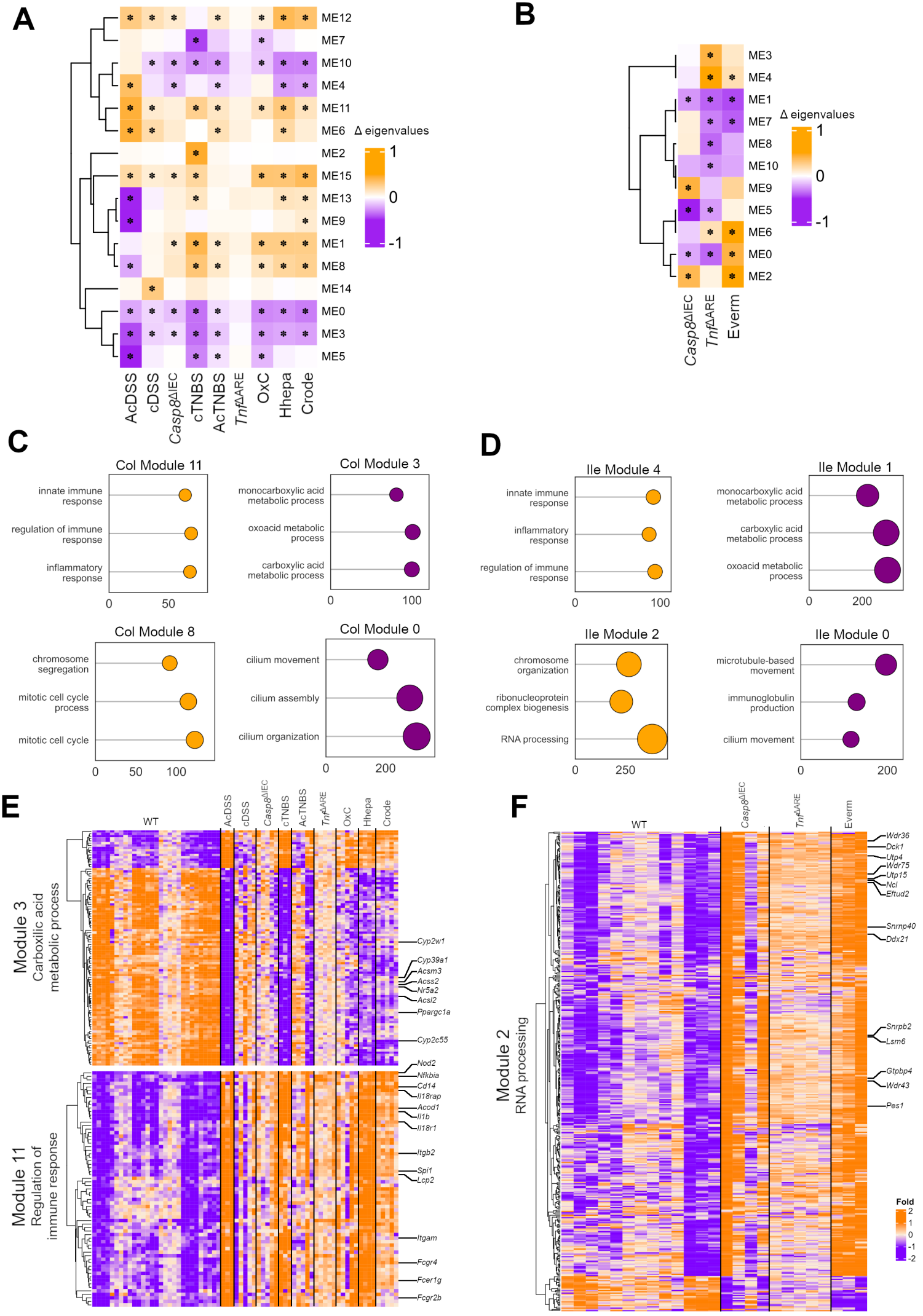

A
ME12
ME7
ME10
ME4
ME11
ME6
ME2
ME15
ME13
ME9
ME1
ME8
ME14
ME0
ME3
ME5
AcDSS
cDSS
Casp8ΔIEC
cTNBS
AcTNBS
TnfΔARE
OxC
Hhepa
Crode
Δ eigenvalues
1
0
-1
B
ME3
ME4
ME1
ME7
ME8
ME10
ME9
ME5
ME6
ME0
ME2
Casp8ΔIEC
TnfΔARE
Everm
Δ eigenvalues
1
0
-1
C
Col Module 11
innate immune response
regulation of immune response
inflammatory response
0
50
Col Module 3
monocarboxylic acid metabolic process
oxoacid metabolic process
carboxylic acid metabolic process
0
100
Col Module 8
chromosome segregation
mitotic cell cycle process
mitotic cell cycle
0
50
100
Col Module 0
cilium movement
cilium assembly
cilium organization
0
200
D
Ile Module 4
innate immune response
inflammatory response
regulation of immune response
0
100
Ile Module 1
monocarboxylic acid metabolic process
carboxylic acid metabolic process
oxoacid metabolic process
0
200
Ile Module 2
chromosome organization
ribonucleoprotein complex biogenesis
RNA processing
0
250
Ile Module 0
microtubule-based movement
immunoglobulin production
cilium movement
0
200
E
WT
AcDSS
cDSS
Casp8ΔIEC
cTNBS
AcTNBS
TnfΔARE
OxC
Hhepa
Crode
Module 3
Carboxilic acid metabolic process
Cyp2w1
Cyp39a1
Acsm3
Acss2
Nr5a2
Acsl2
Ppargc1a
Cyp2c55
Module 11
Regulation of immune response
Nod2
Nfkbia
Cd14
Il18rap
Acod1
Il1b
Il18r1
Itgb2
Spi1
Lcp2
Itgam
Fcgr4
Fcer1g
Fcgr2b
F
WT
Casp8ΔIEC
TnfΔARE
Everm
Module 2
RNA processing
Wdr36
Dck1
Utp4
Wdr75
Utp15
Ncl
Eftud2
Snrnp40
Ddx21
Snrpb2
Lsm6
Gtpbp4
Wdr43
Pes1
Fold
2
1
0
-1
-2

Figure 5

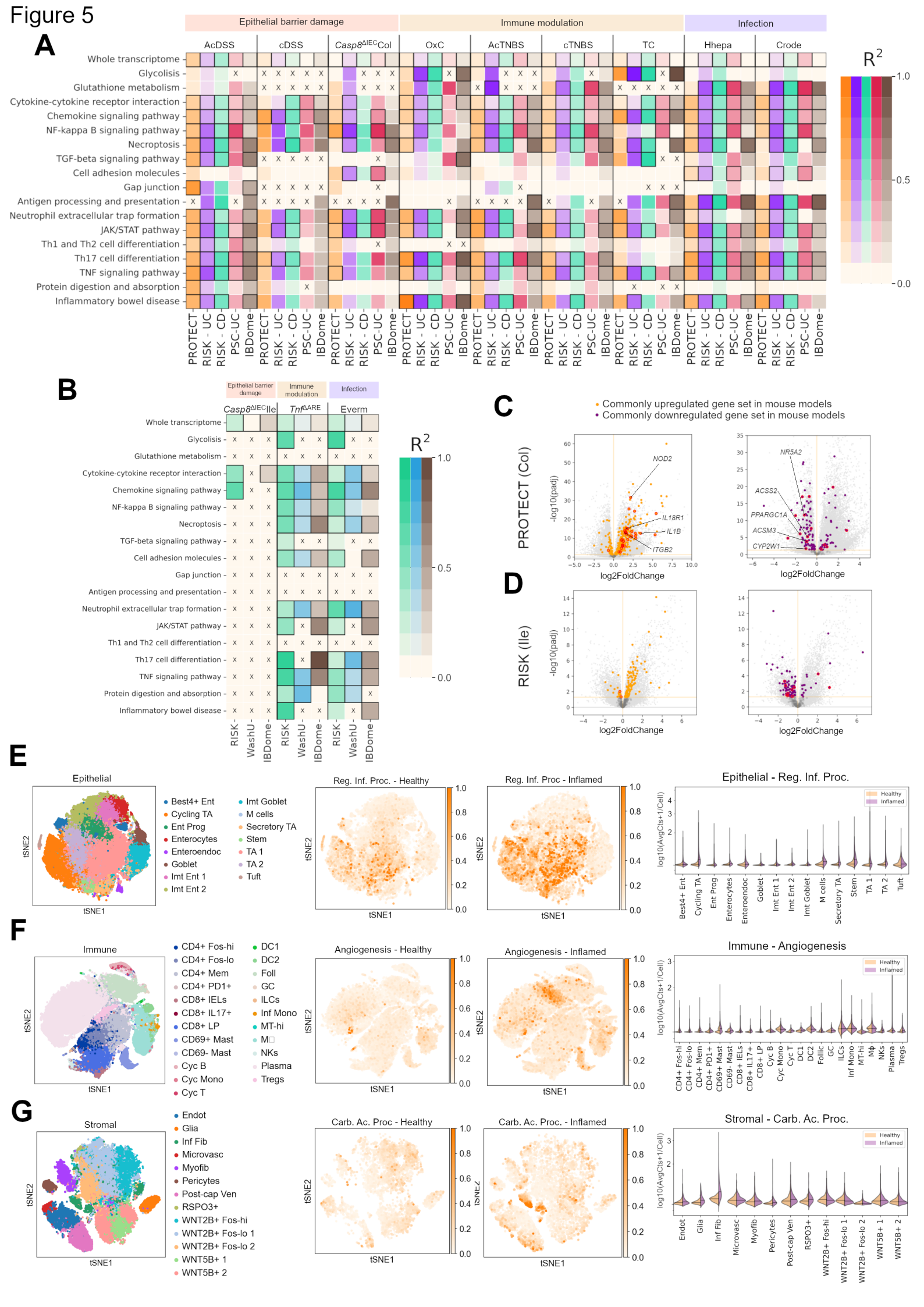
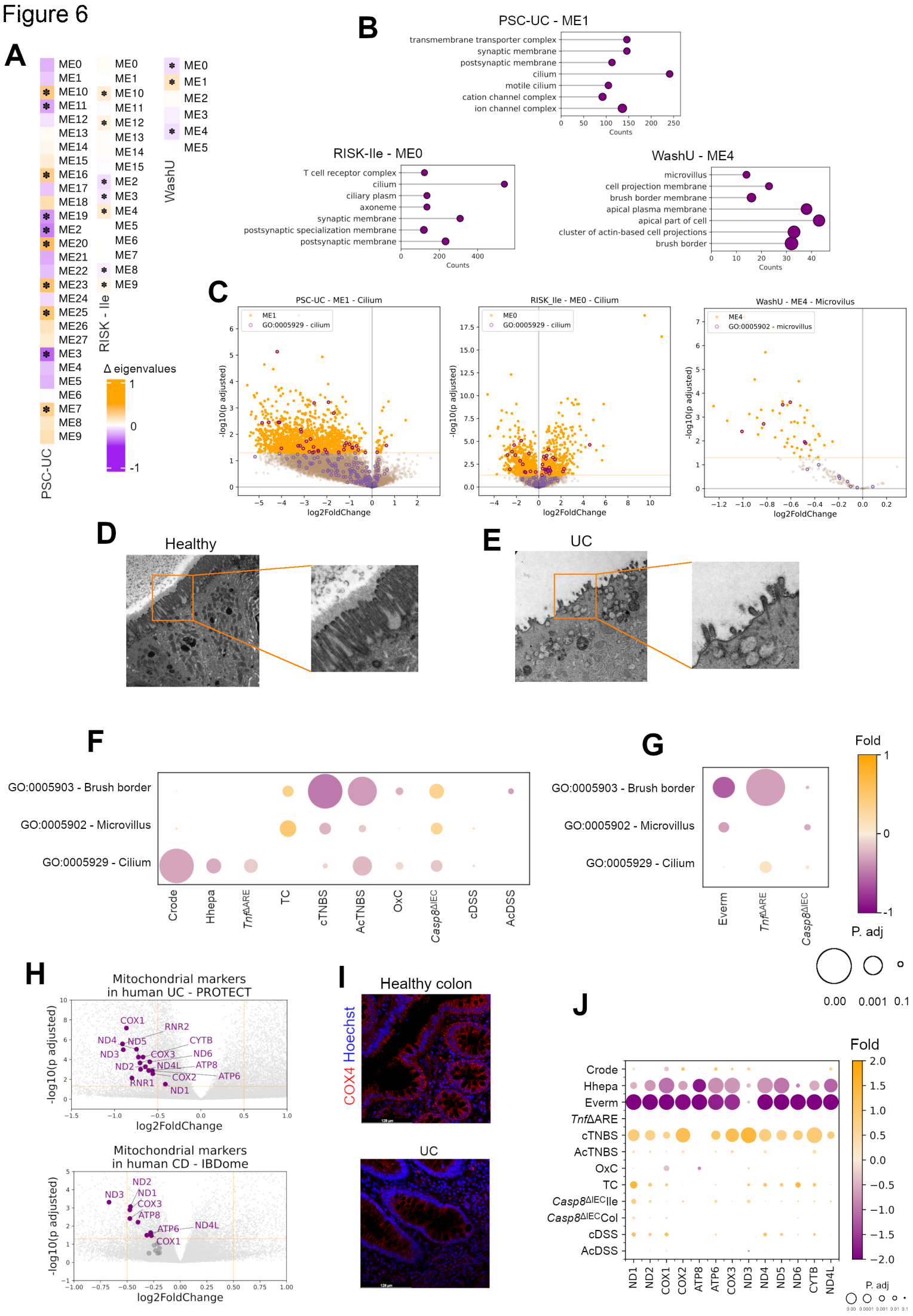
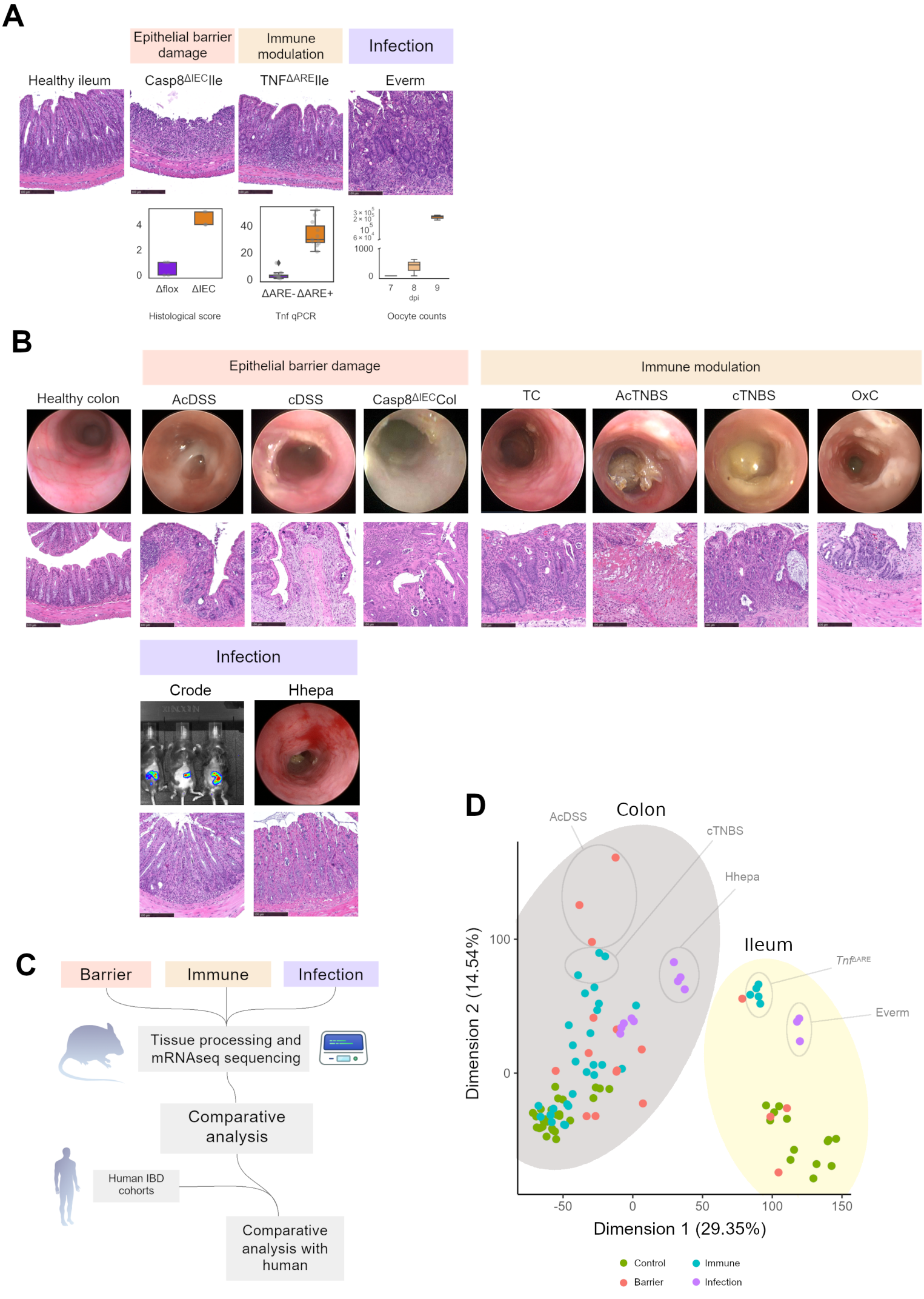

A
Epithelial barrier damage
Immune modulation
Infection
AcDSS
cDSS
Casp8ΔIECCol
OxC
AcTNBS
cTNBS
TC
Hhepa
Crode
Whole transcriptome
Glycolisis
Glutathione metabolism
Cytokine-cytokine receptor interaction
Chemokine signaling pathway
NF-kappa B signaling pathway
Necroptosis
TGF-beta signaling pathway
Cell adhesion molecules
Gap junction
Antigen processing and presentation
Neutrophil extracellular trap formation
JAK/STAT pathway
Th1 and Th2 cell differentiation
Th17 cell differentiation
TNF signaling pathway
Protein digestion and absorption
Inflammatory bowel disease
PROTECT
RISK - UC
RISK - CD
PSC-UC
IBDome
R2
1.0
0.5
0.0
B
Casp8ΔIECIle
TnfΔARE
Everm
RISK
WashU
IBDome
C
Commonly upregulated gene set in mouse models
Commonly downregulated gene set in mouse models
PROTECT (Col)
-log10(padj)
log2FoldChange
NOD2
IL18R1
IL1B
ITGB2
NR5A2
ACSS2
PPARGC1A
ACSM3
CYP2W1
D
RISK (Ile)
E
Epithelial
Best4+ Ent
Cycling TA
Ent Prog
Enterocytes
Enteroendoc
Goblet
Imt Ent 1
Imt Ent 2
Imt Goblet
M cells
Secretory TA
Stem
TA 1
TA 2
Tuft
tSNE1
tSNE2
Reg. Inf. Proc. - Healthy
Reg. Inf. Proc - Inflamed
Epithelial - Reg. Inf. Proc.
log10(AvgCts+1/Cell)
Healthy
Inflamed
F
Immune
CD4+ Fos-hi
CD4+ Fos-lo
CD4+ Mem
CD4+ PD1+
CD8+ IELs
CD8+ IL17+
CD8+ LP
CD69+ Mast
CD69- Mast
Cyc B
Cyc Mono
Cyc T
DC1
DC2
Foll
GC
ILCs
Inf Mono
MT-hi
NKs
Plasma
Tregs
Angiogenesis - Healthy
Angiogenesis - Inflamed
Immune - Angiogenesis
G
Stromal
Endot
Glia
Inf Fib
Microvasc
Myofib
Pericytes
Post-cap Ven
RSPO3+
WNT2B+ Fos-hi
WNT2B+ Fos-lo 1
WNT2B+ Fos-lo 2
WNT5B+ 1
WNT5B+ 2
Carb. Ac. Proc - Healthy
Carb. Ac. Proc. - Inflamed
Stromal - Carb. Ac. Proc.

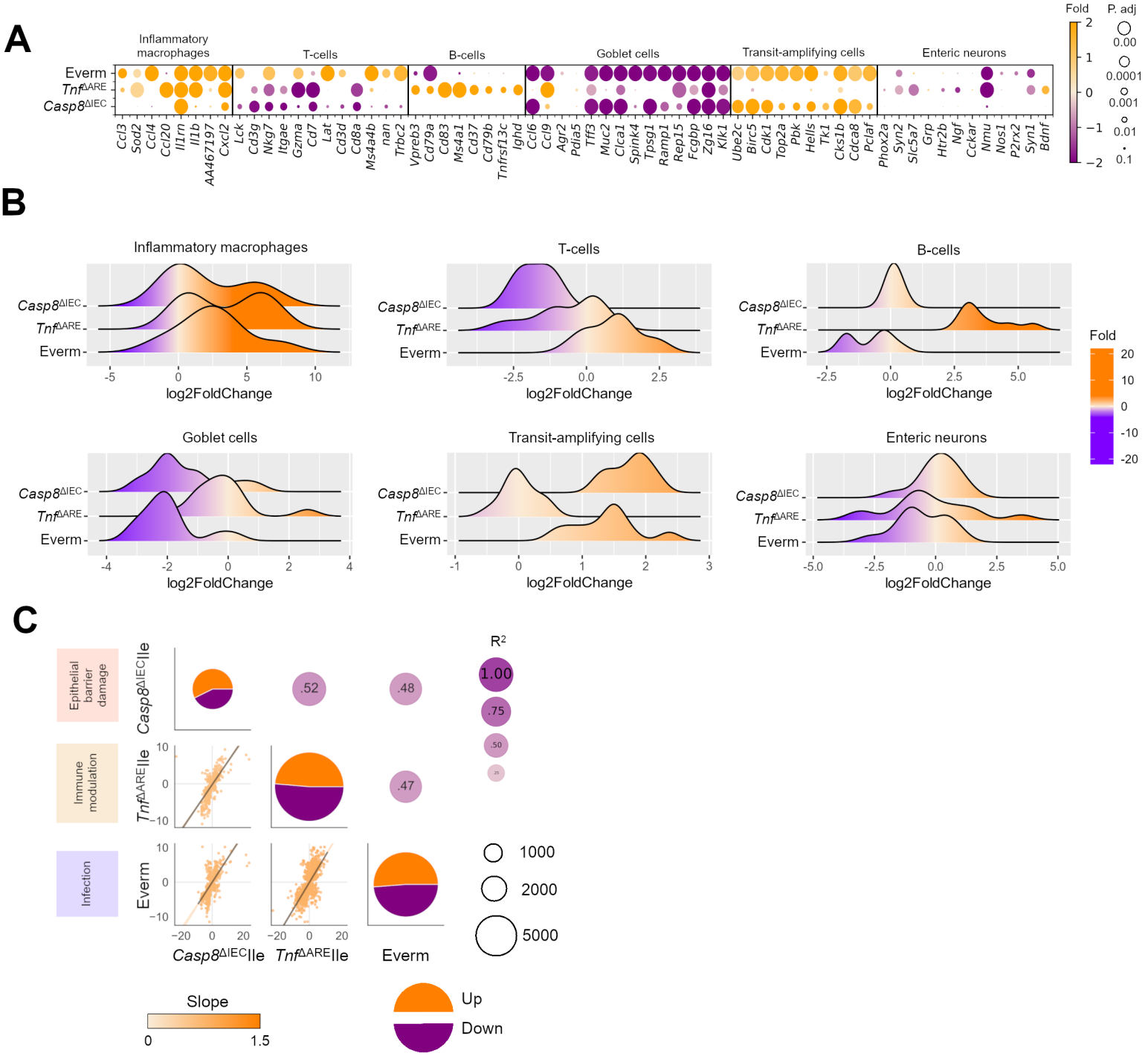
Figure 6


A
PSC-UC
RISK - Ile
WashU
Δ eigenvalues
B
PSC-UC - ME1
transmembrane transporter complex
synaptic membrane
postsynaptic membrane
cilium
motile cilium
cation channel complex
ion channel complex
Counts
RISK-Ile - ME0
T cell receptor complex
cilium
ciliary plasm
axoneme
synaptic membrane
postsynaptic specialization membrane
postsynaptic membrane
WashU - ME4
microvillus
cell projection membrane
brush border membrane
apical plasma membrane
apical part of cell
cluster of actin-based cell projections
brush border
C
PSC-UC - ME1 - Cilium
RISK_Ile - ME0 - Cilium
WashU - ME4 - Microvilus
-log10(p adjusted)
log2FoldChange
D
Healthy
E
UC
F
GO:0005903 - Brush border
GO:0005902 - Microvillus
GO:0005929 - Cilium
Crode
Hhepa
TnfΔARE
TC
cTNBS
AcTNBS
OxC
Casp8ΔIEC
cDSS
AcDSS
G
Everm
Fold
P. adj
H
Mitochondrial markers in human UC - PROTECT
Mitochondrial markers in human CD - IBDome
I
Healthy colon
COX4 Hoechst
UC
J
ND1
ND2
COX1
COX2
ATP8
ATP6
COX3
ND3
ND4
ND5
ND6
CYTB
ND4L

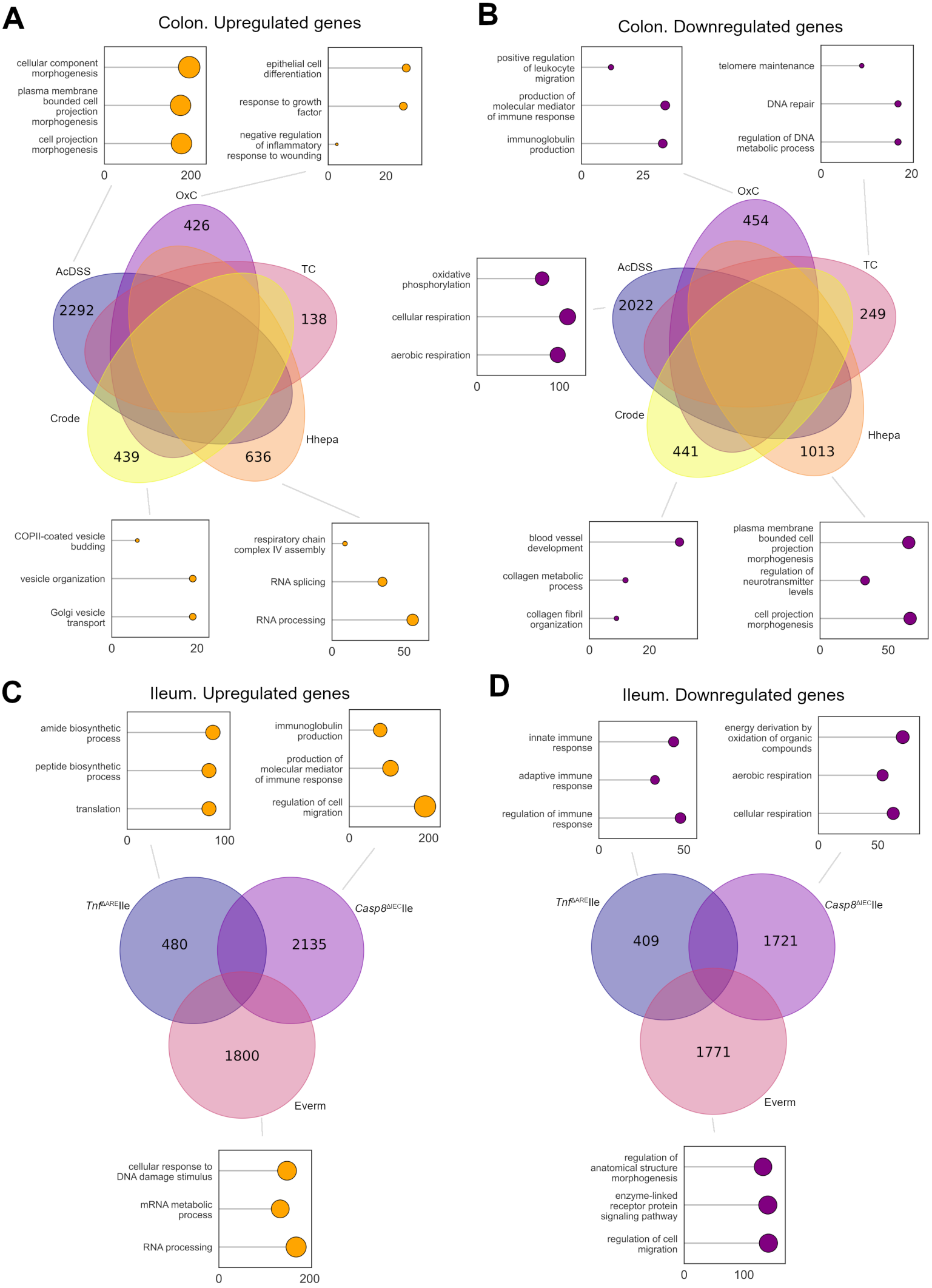
Supplementary figure 1



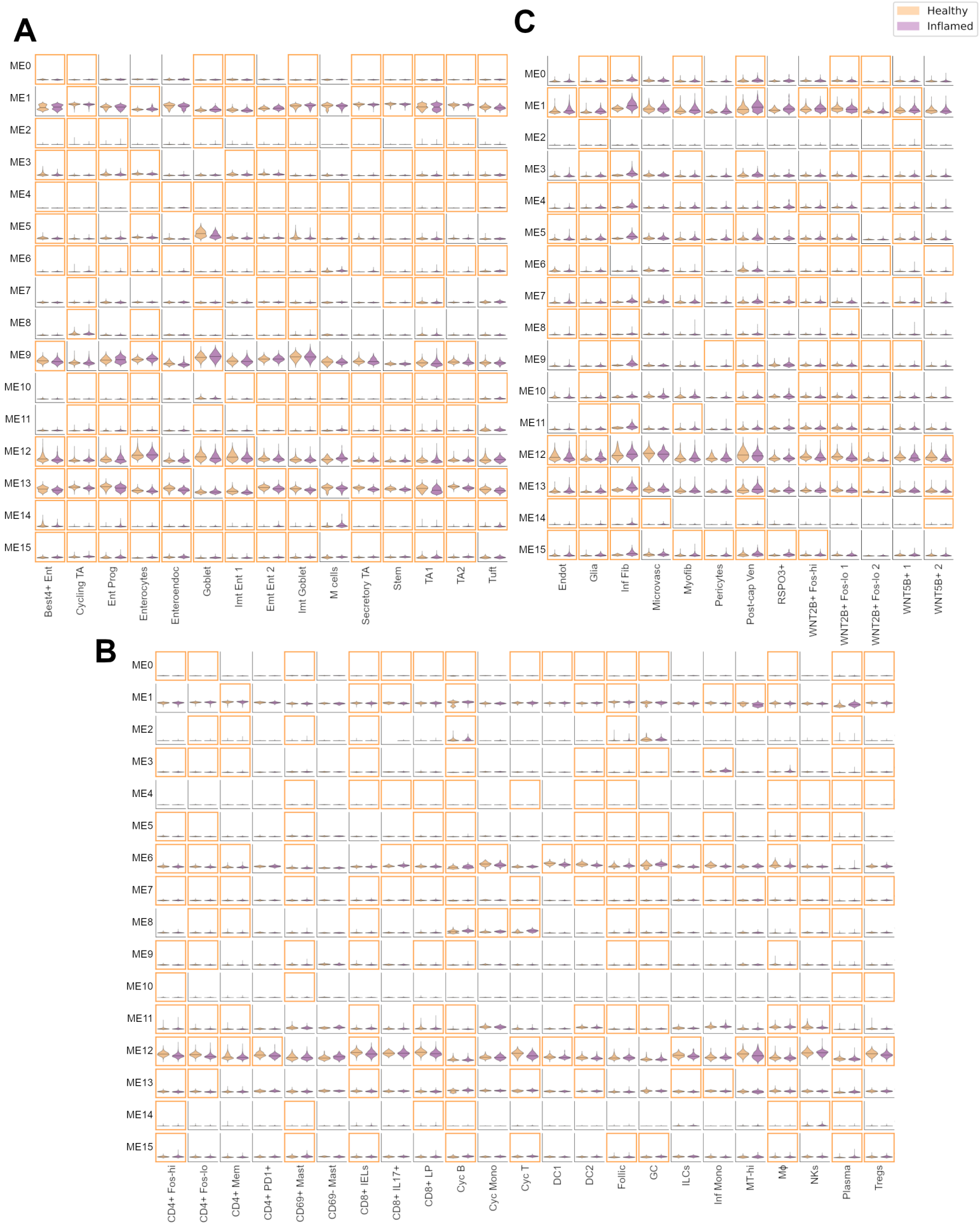
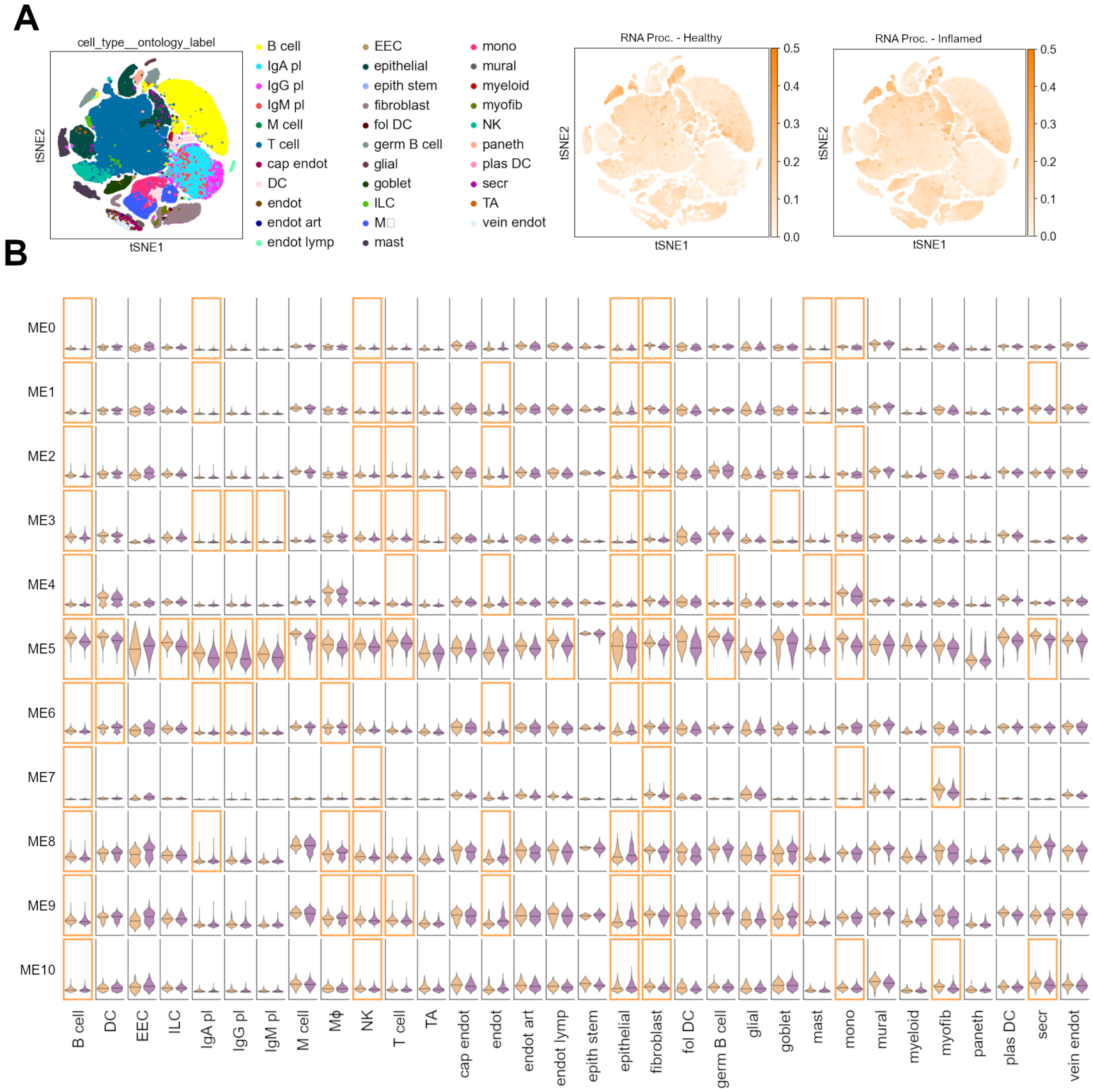
Supplementary figure 2



Supplementary figure 3



Supplementary figure 4



Supplementary figure 5