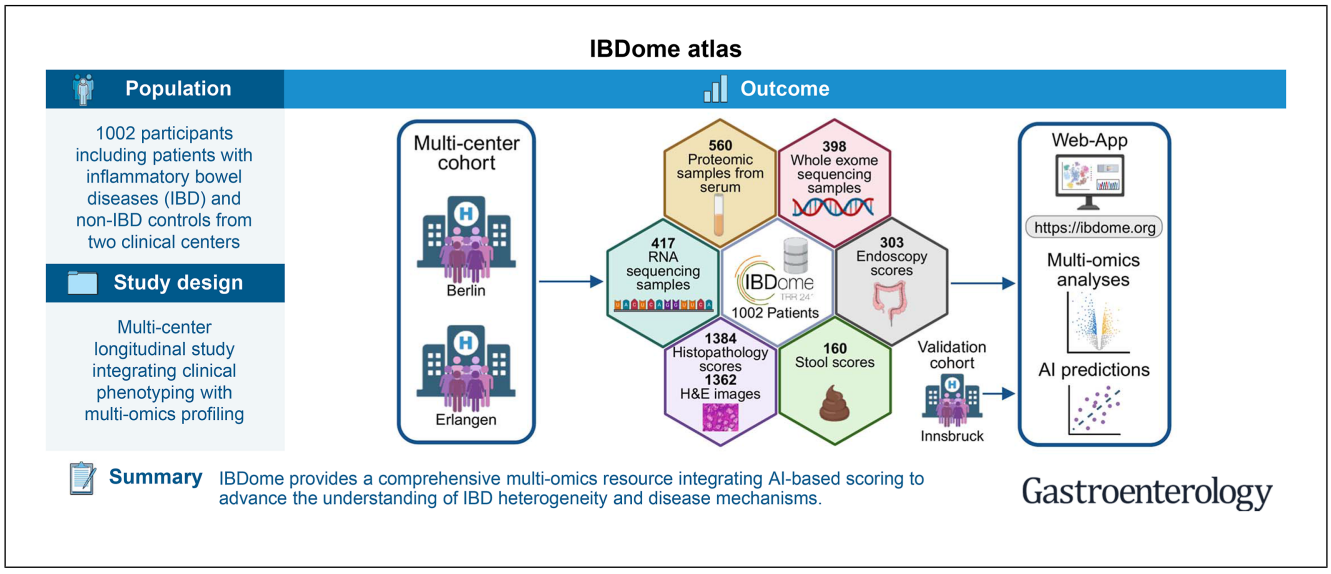
Gastroenterology 2026;171:902–918

IBDome: An Integrated Molecular, Histopathological, and Clinical Atlas of Inflammatory Bowel Diseases

Christina Plattner,[1,*] **Gregor Sturm**,[1,2,*] **Anja A. Kühl**,[3,*] **Raja Atreya**,[4,5,*] **Sandro Carollo**,[1,*] Dietmar Rieder,[1] Raphael Gronauer,[1] Michael Günther,[6] Steffen Ormanns,[6] Claudia Manzl,[7] Stefan Wirtz,[4] Asier Rabasco Meneghetti,[8] Ahmed N. Hegazy,[9,10,11] Jay V. Patankar,[4,5] Zunamys I. Carrero,[8] Felix Grabherr,[12] Moritz Meyer,[12] TRR241 IBDome Consortium, Timon E. Adolph,[12] Herbert Tilg,[12] Markus F. Neurath,[4,5] Jakob Nikolas Kather,[8,13,14,15] **Christoph Becker**,[4,5,§] **Britta Siegmund**,[9,11,§] and **Zlatko Trajanoski**[1,§]

[1]Biocenter, Institute of Bioinformatics, Medical University of Innsbruck, Innsbruck, Austria; [2]Boehringer Ingelheim International Pharma GmbH & Co KG, Biberach, Germany; [3]iPATH.Berlin, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany; [4]Department of Medicine 1, Friedrich-Alexander-Universität Erlangen-Nürnberg and Universitätsklinikum Erlangen, Erlangen, Germany; [5]Deutsches Zentrum Immuntherapie, Erlangen, Germany; [6]Institute of General Pathology, Medical University of Innsbruck, Innsbruck, Austria; [7]Institute of Pathology, Neuropathology and Molecular Pathology, Medical University of Innsbruck, Innsbruck, Austria; [8]Else Kroener Fresenius Center for Digital Health, Faculty of Medicine, and University Hospital Carl Gustav Carus, Dresden University of Technology, Dresden, Germany; [9]Medizinische Klinik für Gastroenterologie, Infektiologie und Rheumatologie, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; [10]Deutsches Rheuma-Forschungszentrum, a Leibniz Institute, Berlin, Germany; [11]Cluster of Excellence ImmunoPreCept, Charité–Universitätsmedizin Berlin, Berlin, Germany; [12]Department of Internal Medicine I, Gastroenterology, Hepatology, Endocrinology and Metabolism, Medical University of Innsbruck, Innsbruck, Austria; [13]Department of Medicine I, Faculty of Medicine, and University Hospital Carl Gustav Carus, Dresden University of Technology, Dresden, Germany; [14]Medical Oncology, National Center for Tumor Disease, University Hospital Heidelberg, Heidelberg, Germany; and [15]Pathology and Data Analytics, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds, United Kingdom



BACKGROUND & AIMS: Multiomic and multimodal datasets with detailed clinical annotations offer significant potential to advance understanding of inflammatory bowel diseases (IBDs), refine diagnostics, and enable personalized therapeutic strategies. **METHODS:** In this multicohort study, an extensive multiomic and multimodal analysis of 1002 clinically annotated patients with IBD and non-IBD controls was performed, incorporating whole-exome and RNA sequencing of normal and inflamed gut tissues, serum proteomics, and histopathological assessments from images of H&E-stained tissue sections. **RESULTS:** Transcriptomic profiles of normal and inflamed tissues revealed distinct site-specific inflammatory signatures in Crohn's disease and ulcerative colitis. Leveraging serum proteomics, an inflammatory protein severity signature that reflects underlying intestinal molecular inflammation was developed. Furthermore, foundation model-based deep learning accurately predicted histologic disease activity scores and enabled Crohn's disease vs ulcerative colitis classification from images of H&E-

stained intestinal tissue sections, offering a robust tool for clinical evaluation. **CONCLUSIONS:** This integrative, publicly available multiomics resource for IBD research highlights the potential of combining multiomics and advanced computational approaches to improve understanding and management of IBD.

Keywords: Crohn's Disease; Ulcerative Colitis; Multiomics Database; Artificial Intelligence.

Inflammatory bowel diseases (IBD) manifest as 2 major subtypes, ulcerative colitis (UC) and Crohn's disease (CD). In UC, the inflammation is limited to the mucosa and submucosa of the colon and continuously spreads to a varying extent from the rectum to the proximal colon and, in severe cases, to the terminal ileum, whereas CD affects all layers of the gastrointestinal wall. Symptoms of both UC and CD include diarrhea, rectal bleeding, abdominal pain, weight loss, and fatigue.[1] IBD increases the risk of colorectal cancer[2] and of concomitant manifestation of other immune-mediated inflammatory conditions, such as arthritis.[3] The disease affected 4.9 million people worldwide in 2019, and both incidence and prevalence have been increasing globally since 1990.[4] The exact cause of the disease is not known, but the leading hypothesis is that it arises from a combination of genetic predisposition, dysbiosis of the gut microbiome, and environmental factors leading to excessive activation of the mucosal immune system.[5]

Owing to the growing understanding of the immunopathogenesis of IBD, several classes of biological therapies (targeting tumor necrosis factor [TNF], interleukin [IL] 12, IL23, IL23p19, and $\alpha 4\beta 7$ integrins), and small molecules (JAK inhibitors and S1P receptor modulators) have been approved for treatment.[6,7] The advent of these targeted therapies has substantially improved the therapeutic outcomes of patients and has made a major impact on existing therapeutic algorithms. With this progress in advanced therapies, treatment goals have evolved simultaneously. Traditionally, treatment was aimed mainly at symptom control. Currently, restoration of quality of life, as well as biochemical, endoscopic, and even histologic remission are strived for, with the potential to modulate the subsequent disease course and prevent disease-associated complications (eg, hospitalization, penetrating disease, colorectal neoplasia, and surgery). Furthermore, novel endpoints that could also have an impact on the subsequent disease course, such as histologic remission and transmural healing, are being strived for.[8] Nonetheless, these therapies are not yet optimally used because response to therapy is highly heterogeneous among patients. Depending on the outcome parameter and treatment duration, more than one-half of advanced therapy–treated patients do not demonstrate adequate response to therapy or are prone to lose response to therapy in the course of treatment. An apparent deficit to fully comprehend the molecular pathologic processes in IBD has hindered us from determining the most appropriate therapy based on the underlying biologic processes they target, resulting in increased likelihood of failure.[9]

WHAT YOU NEED TO KNOW

BACKGROUND AND CONTEXT

Inflammatory bowel disease is complex and heterogeneous with no single reliable standard for diagnosis or activity assessment, highlighting the need for integrated, objective approaches combining clinical, molecular, and tissue information.

NEW FINDINGS

IBDome integrates data from more than 1000 patients, reveals site-specific inflammatory patterns, develops a blood-based protein signature for noninvasive disease activity monitoring, and shows that artificial intelligence can assess tissue inflammation.

LIMITATIONS

This study is limited due to the use of a predefined 96-plex serum proteomic panel and a cross-sectional design, which precludes assessment of temporal dynamics and outcome prediction. In addition, the non-IBD control group was moderately older than the IBD cohort.

CLINICAL RESEARCH RELEVANCE

IBDome links clinical data with blood and tissue measures of inflammation, offering a blood-based protein signature for noninvasive disease monitoring and artificial intelligence analysis to reduce variability in histologic assessment. These integrated tools support more objective evaluation of disease activity and may help guide treatment decisions and monitoring strategies in inflammatory bowel disease.

BASIC RESEARCH RELEVANCE

IBDome is a publicly accessible, integrated resource enabling researchers to study inflammatory bowel disease across genetic, molecular, tissue, and clinical levels. By combining multilayer data with artificial intelligence applied to tissue images, the atlas supports investigation of disease mechanisms, site-specific inflammation, and development of computational approaches for objective tissue-based disease assessment.

Clinical symptoms do not always reliably reflect disease activity, as patients may experience significant inflammation without overt symptoms or report severe symptoms

* Authors share co-first authorship; § Authors share co-senior authorship.

Abbreviations used in this paper: **bMIS, biopsy molecular inflammation score; CCL, CC-chemokine ligand; CD, Crohn's disease; HBI, Harvey-Bradshaw Index; HGF, hepatocyte growth factor; IBD, inflammatory bowel disease; IBD-IPSS, Inflammatory Bowel Disease Inflammatory Protein Severity Signature; IFN, interferon; IL, interleukin; MMP, metalloproteinase; MUC, mucin; OSM, oncostatin M; PMS, Partial Mayo Score; RNA-seq, RNA sequencing; SES-CD, Simple Endoscopic Score for Crohn's Disease; TGF, transforming growth factor; TNF, tumor necrosis factor; UC, ulcerative colitis; UCEIS, Ulcerative Colitis Endoscopic Index of Severity; WSI, whole-slide image.**

Most current article

© 2026 The Author(s). Published by Elsevier Inc. on behalf of American Gastroenterological Association Institute. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

0016-5085
https://doi.org/10.1053/j.gastro.2026.05.023

despite minimal inflammatory activity. This inconsistency underscores the need for objective measures of disease activity to guide clinical decision making and improve patient outcomes. However, there is no single "gold standard" for diagnosing IBD, assessing disease severity, or evaluating treatment response. A multifaceted approach used by physicians integrates clinical symptoms, laboratory biomarkers, radiologic imaging, endoscopic examinations, and histologic analysis of biopsy specimens.[10] Although this comprehensive strategy provides valuable insights, it also highlights the complexities of assessing disease activity and the ongoing need for standardized, objective, and accessible diagnostic tools.

Several large-scale efforts have laid the foundation for integrative IBD research, including the Human Microbiome Project Phase 2[11,12] and the 1000IBD project,[13] which have offered critical insights into microbial dynamics, host–microbe interactions, and inflammatory pathways. Similarly, the IBD Transcriptome and Metatranscriptome Meta-Analysis[14] and other regional biobanks have integrated valuable transcriptomic and clinical data from publicly available datasets. However, many of these efforts focus on a subset of data modalities, prioritize either longitudinal or cross-sectional designs, or lack systematic integration of histopathologic features at scale. In parallel, single-cell atlases of IBD, including studies by Kong et al[15] and Smillie et al,[16] have revealed disease-relevant immune cell states at high resolution, but are often restricted in sample size or clinical diversity.

In this study, we addressed these challenges by creating a comprehensive multicenter, multiomic, and multimodal IBD atlas (IBDome atlas), integrating individual genomic, transcriptomic, proteomic, histopathologic, and clinical data from 1002 patients with IBD and respective controls. Using this resource, we investigated site-specific immunologic pathways and features, developed a novel serum protein-based disease activity signature (IBD- Inflammatory Protein Severity Signature [IPSS]), and leveraged deep learning prediction of histologic disease activity from histologic images using general-purpose foundation models.

Materials and Methods

The materials and methods of this study are described in detail in the Supplementary Material.

Study Design and Cohorts

This study was designed as a multicenter, multimodal analysis of IBD integrating clinical, endoscopic, histopathologic, molecular, and imaging data. Patients with CD, UC, and non-IBD controls were recruited at tertiary referral centers in Berlin and Erlangen, Germany. The IBDome study centers are located at the Department of Medicine 1, Universitätsklinikum Erlangen and the Department of Gastroenterology, Infectious Diseases and Rheumatology including Clinical Nutrition at the Charité–Universitätsmedizin Berlin. An independent cohort from Innsbruck, Austria was included for external validation of selected analyses. All participants provided written informed consent, and the study was approved by the respective institutional ethics committees. Sample collection, processing, and data generation followed harmonized protocols across centers.

Clinical and Endoscopic Assessment

A standardized and unified medical questionnaire was designed and implemented as part of the clinical information systems of both study centers. The questionnaire consists of the following 2 parts: (1) basic data, which was entered at the initial visit, including birth year, sex, diagnosis, and pre-existing conditions; and (2) time course longitudinally collected data, which the attending physician enters at each visit, including body weight, disease activity scores, and ongoing medication. Clinical disease activity is recorded as Partial Mayo Score[17] for UC and Harvey-Bradshaw Index[18] for CD.

Endoscopic disease activity was assessed using validated scoring systems, including the Ulcerative Colitis Endoscopic Index of Severity (UCEIS)[19] for UC and the Simple Endoscopic Score for Crohn's Disease (SES-CD).[20] The scoring was done based on the established criteria of both scores by experienced endoscopists at both participating centers. The endoscopists were blinded to the individual molecular date of the investigated patients.

Clinical and endoscopic data were curated, standardized, and harmonized across cohorts before downstream analyses.

Histopathology and Disease Activity Scoring

Formalin-fixed paraffin-embedded intestinal tissue samples were processed using routine clinical protocols and stained with H&E. Histologic disease activity was assessed based on established scoring systems, Riley[21] for UC and Naini Cortina[22] for CD. For the Berlin and Erlangen cohort, histologic scoring was performed by a gastrointestinal histology expert following standardized criteria. For the independent Innsbruck cohort, histologic disease activity was assessed by 2 gastrointestinal histology experts and consensus scores were derived by averaging the independent assessments.

To ensure comparability between different histopathology scores (modified Naini Cortina Score and modified Riley[21] score), we normalized the scores to a 0–1 scale, considering the tissue-specific maximum score for each disease entity (CD or UC) and sampling method (biopsy or resection). The maximum scores are listed in Table 1.

RNA-Sequencing Library Preparation and Sequencing

Biopsies collected during endoscopy or from resected tissue using a single-use biopsy forceps (Olympus) were incubated in RNAprotect Tissue Reagent (Qiagen) and stored at −80°C. For RNA isolation, biopsies were thawed on ice and homogenized in RLT buffer (Qiagen) using the TissueLyser LT (Qiagen). RNA was isolated, cleaned, and concentrated using the RNeasy kit (Qiagen) and RNA Clean & Concentrator kit (Zymo Research). The concentration was measured at NanoDrop One/One (Thermo Fisher Scientific) and quality (RNA integrity number) at TapeStation (Agilent). RNA was shipped on dry ice to the NGS Competence Center Tübingen for sequencing.

Table 1. Maximum Histopathology Scores for the Modified Naini Cortina[22] and Modified Riley[21] Scores Categorized by Tissue Type and Sampling Method (Biopsy or Resection)

Tissues and sampling method	Maximum modified Naini Cortina score	Maximum modified Riley score
Colon, rectum, and cecum		
Resection	20	21
Biopsy	16	17
Ileum, ileocecal valve, small intestine, and anastomosis, pouch		
Resection	14	16
Biopsy	10	12

Whole-Exome Sequencing Library Preparation and Sequencing

Total DNA was isolated from whole blood using the DNeasy Blood & Tissue Kit according to the manufacturer's protocol (Qiagen). The concentration was measured using NanoDrop One/One (Thermo Fisher Scientific). The DNA was shipped on dry ice to the NGS Competence Center, Tübingen, Germany for sequencing.

Proteomic Profiling

A serum sample aliquot was thawed on ice for 1 hour and centrifuged at 3000 rpm for 1 minute at 4°C. 96-well full skirted polymerase chain reaction plates (catalog number AB0800; ThermoFisher Scientific) were used with 80 μL of serum per well and sealed with adhesive tape (MicroAmp seal, catalog number 4306311; ThermoFisher Scientific). The pipetting scheme for all plates was randomized by the BIH Core Unit Proteomics. Samples were shipped on dry ice to the BIH Core Unit Proteomics, Charité, Berlin for measurements with the Olink Target 96 Inflammation panel. Protein abundance measurements were quality-controlled, normalized, and batch-corrected before analysis.

Whole-Slide Imaging and Preprocessing

Whole-slide images (WSIs) of H&E-stained tissue sections were scanned in 2 batches at different centers: Medical University of Innsbruck (Innsbruck) and Charité (Berlin). WSIs from the first batch were scanned at ×40 magnification using a NanoZoomer S210 slide scanner (Hamamatsu), and the analysis was performed using NDP.view2 software (Hamamatsu). WSIs from the second batch were scanned at ×100 magnification using a Vectra3 automated quantitative pathology imaging system (Akoya Biosciences).

The images underwent quality control to exclude artifacts and noninformative regions. Tissue regions were segmented and subdivided into nonoverlapping tiles for downstream analyses. Standardized preprocessing steps, including color normalization[23] and artifact filtering, were applied to mitigate technical variability.

Image-Derived Feature Extraction and Deep Learning Analyses

Image-derived feature representations were extracted from H&E-stained tissue sections using pretrained foundation models for histopathology[24–27] and aggregated into slide-level embeddings for downstream analysis. Attention-based multiple instance learning frameworks[28] were used for weakly supervised regression and classification tasks, including disease subtype classification and histologic disease activity prediction. For the latter, models were trained using cross-validation on the Berlin cohort and deployed to the Erlangen and Innsbruck cohorts, with predictions aggregated across folds. Disease subtype classification followed a train-deployment paradigm, with training on the Berlin cohort and evaluation on the Erlangen cohort. Model performance was assessed against expert-assigned labels or consensus histologic scores, as appropriate, and stratified analyses were performed on predefined tissue subsets. Model interpretability was evaluated using attention heatmaps highlighting regions contributing most strongly to predictions.

Results

Development of the IBDome Atlas

We first generated multiomic and multimodal data, encompassing clinical metadata from 1002 patients with IBD and non-IBD controls, including histopathology, high-resolution H&E images, whole-exome sequencing, RNA sequencing (RNA-seq), serum proteomics data, endoscopic activity scores, stool appearance scores, and clinical disease characteristics to comprehensively characterize the underlying immunopathogenesis of IBD (Figure 1*A* and *B* and Supplementary Figure 1*A*). In total, this atlas includes data from 539 patients diagnosed with CD, 321 patients diagnosed with UC, 26 patients diagnosed with indeterminate colitis, and 116 non-IBD controls without any intestinal inflammatory condition from 2 distinct study centers in Berlin and Erlangen, Germany (Figure 1*C*). Longitudinal clinical information is available for almost all patients with IBD, and for a subset, repeated samples were collected over time. To facilitate exploration of the clinical and molecular data, we developed an interactive and publicly available

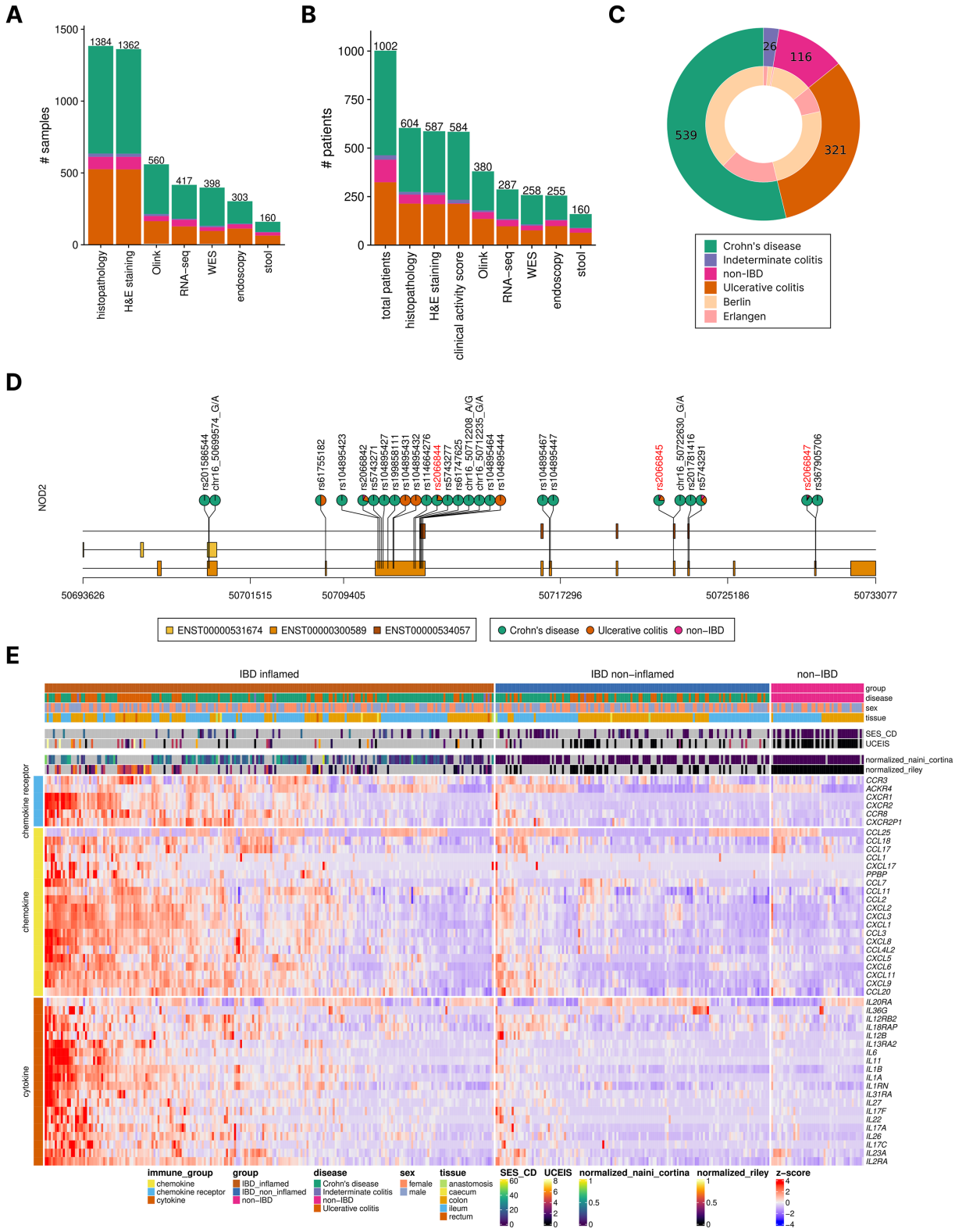


Figure 1. Characteristics of the IBDome atlas. (*A*) Schematic overview of the datasets and sample numbers for the 1002 patients integrated in IBDome. (*B*) Number of patients per sample type; *colors* represent the different diseases and *numbers* on top of the *graphs* depict the total numbers. (*C*) Patient distribution is illustrated as a *nested pie chart*, with the *outer circle* representing the number of patients per disease and the *inner circle* indicating the proportion of patients per study center (Berlin and Erlangen, Germany). (*D*) *Exome mutation map* of NOD2; the known most frequent variants are highlighted in *red*. (*E*) *Heatmap* of differentially expressed cytokines, chemokines, and chemokine receptors between IBD inflamed samples (n = 223) and non-IBD controls (n = 46) clustered by Euclidean distance and complete linkage.

web application denoted IBDome Explorer (https://ibdome.org).

Genomic and transcriptomic characterization supports that the IBDome atlas captures key molecular features of IBD (Figure 1*D* and *E*). As expected, mutations in NOD2 are observed predominantly in patients with CD. The 3 most common variants (ie, R702W, G908R, and 1007fs)[29] exhibit higher mutation frequencies compared with UC and non-IBD patients (Figure 1*D*). To assess the genetic regulation of tissue gene expression, we performed expression quantitative trait locus analysis (Supplementary Figure 1*B* and *C*, Supplementary Tables 1–5, Supplementary Results).

Differential expression analysis between inflamed IBD (ie, tissue- and date-matching histopathology score > 0) and non-IBD control samples showed an up-regulation of cytokines, chemokines, and chemokine receptors associated with disease severity scores determined by histopathology or endoscopy scores (Figure 1*E*, Supplementary Table 6). To further characterize the immune landscape, we performed immune cell deconvolution using the quanTIseq method,[30] which indicated an increased abundance of B cells, alongside a reduced fraction of M2 macrophages in inflamed IBD tissues from both colon and ileum. In contrast, a reduced fraction of regulatory T cells was observed in inflamed colon tissue only (Supplementary Figure 1*D*). Furthermore, disease activity scores (ie, modified Naini Cortina[22] and modified Riley[21] scores evaluated through histopathology, UCEIS,[19] and SES-CD[20] assessed by endoscopy, Bristol stool score, and the clinical activity scores Harvey-Bradshaw Index[18] and Partial Mayo Score[17]) showed significant positive correlations (Supplementary Figure 1*E*–*G*), highlighting their interconnectedness in capturing the severity and progression of IBD.

Multiomic factor analysis revealed tissue- and inflammation-driven variation (Supplementary Figure 2*A*–*D*, Supplementary Results).

Molecular Disease Activity Scoring to Enhance Inflammatory Bowel Disease Assessment

The assessment of disease severity in IBD is crucial for selecting appropriate treatment regimens and adequately assessing response to initiated therapies. Although existing scores demonstrate significant positive correlations with underlying severity of disease (Supplementary Figure 1*E*–*G*), a definitive measure capable of identifying disease activity, including subclinical inflammation that may persist undetected at the molecular level, has yet to be established. Argmann et al[31] recently introduced biopsy- and blood-based molecular signatures—the biopsy molecular inflammation score (bMIS) and the circulating molecular inflammation score—derived from RNA-seq data to evaluate disease severity. Following this approach, we computed the bMIS scores, which effectively distinguished inflamed IBD from noninflamed IBD and non-IBD control groups (Supplementary Figure 3*A*). However, because our cohort lacks blood RNA-seq data, we chose to focus on an alternative approach based on circulating proteins rather than transcriptomic measurements. To this end, we developed the IBD-IPSS, a complementary approach based on the quantification of serum proteins derived from patients' blood. First, we performed principal component analysis for detecting potential confounding factors (Supplementary Figure 3*B*). Subsequently, we used the methodology outlined by Argmann et al[31] to conduct a differential protein abundance analysis comparing samples from actively inflamed and noninflamed patients (Figure 2*A*). For each of the 3 subtypes (ie, IBD, UC, and CD), significantly up-regulated proteins were identified and incorporated into distinct inflammatory protein severity signatures: IBD-IPSS (42 proteins), UC-IPSS (32 proteins), and CD-IPSS (25 proteins), with 17 proteins shared across all conditions (Figure 2*B*, Supplementary Table 7). We then compared these protein-based signatures with the circulating molecular inflammation scores and found that a single protein, namely oncostatin M (OSM),[32] was shared among all signatures (Supplementary Figure 3*C*).

To evaluate the IBD-IPSS, we performed an in silico protein–protein interaction analysis, which indicated that proteins from our signature are predominantly implicated in cytokine-related pathways (Supplementary Figure 3*D*). The protein–protein interaction network analysis identified 5 major clusters, all of which have been determined to be critical processes in the pathophysiology of IBD[33–36]: neutrophil chemotaxis, IL6 family signaling, IL7 signaling, IL18-mediated signaling pathways, and positive regulation of cellular respiration (Figure 2*C*). Because a direct comparison with blood-derived RNA-seq scores is not possible within our cohort, we evaluated the correlation between the computed IPSS score (Supplementary Table 8) and several established inflammatory outcome measures, including endoscopic scores (ie, UCEIS and SES-CD), histopathology scores (ie, modified Riley[21] and modified Naini Cortina[22] scores), clinical activity scores (ie, Partial Mayo Score[17] and Harvey-Bradshaw Index[18]), and computed molecular inflammation scores (ie, bMIS-UC and bMIS-CD) (Supplementary Table 9). The results demonstrated that the serum protein signatures exhibit the strongest correlation with endoscopic scores, with a Pearson correlation coefficient of 0.75 for UC-IPSS and UCEIS and 0.58 for CD-IPSS and SES-CD (Figure 2*D*). To complete the serum protein characterization, we compared inflamed and noninflamed IBD samples with non-IBD controls (Supplementary Figure 3*E*), confirming that OSM levels are significantly elevated during inflammation. Notably, TNF and AXIN1 showed a significant increase in inflamed and noninflamed IBD compared with non-IBD controls, suggesting that these markers may serve as effective biomarkers for IBD, irrespective of the activity status.

Distinct Immunologic Pathways Underpin Site-Specific Inflammatory Signatures in Inflammatory Bowel Disease

In recent years, mounting evidence has highlighted substantial disparities between ileal and colonic CD across diverse intestinal layers. Colonic CD has been observed to manifest disease characteristics comparable with those of UC, reinforcing the notion that IBD encompasses a more

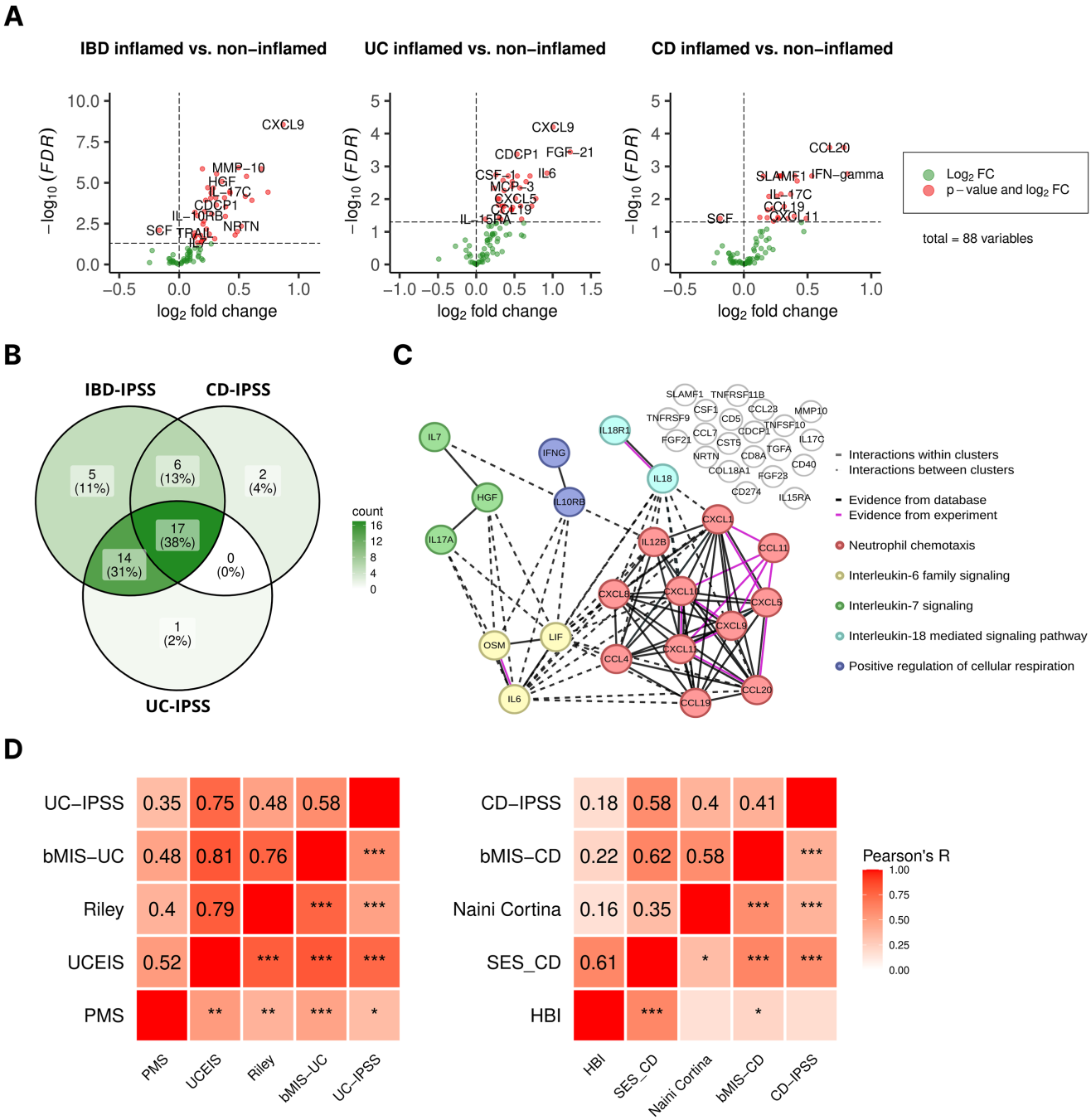


Figure 2. IPSS. (*A*) *Volcano plots* of differentially abundant serum proteins in IBD-inflamed vs noninflamed, UC inflamed vs noninflamed, and CD inflamed vs noninflamed samples assessed by Welch *t* test with an adjusted *P* value $< .1$. FDR, false discovery rate. (*B*) Overlap of proteins in the different inflammatory protein severity signatures. (*C*) Protein–protein interaction network of the serum proteins of the IBD-IPSS. (*D*) Pearson correlation of the inflammatory protein severity signatures with bMIS-UC and bMIS-CD scores derived from gene set variation analysis from RNA-seq data, histopathology scores (ie, normalized modified Riley[21] score, normalized modified Naini-Cortina[22] score), endoscopic scores (ie, UCEIS and SES-CD) and clinical activity scores (ie, Partial Mayo Score[17] [PMS] for UC and Harvey-Bradshaw Index[18] [HBI] for CD). *$P < .05$; **$P < .01$; ***$P < .001$.

intricate spectrum of disease manifestations beyond the conventional classifications of CD and UC.[37,38] The principal component analysis of RNA-seq profiles from the IBDome atlas underscored that tissue type accounted for the largest variance, followed by inflammation grade. Notably, there was no clear visual separation between the overall disease entities CD and UC (Figure 3*A*). Subsequently, we grouped the transcriptomic samples by disease entity, sampling site, and histologic disease activity (ie, CD colon inflamed, CD ileum inflamed, and UC colon inflamed) and performed

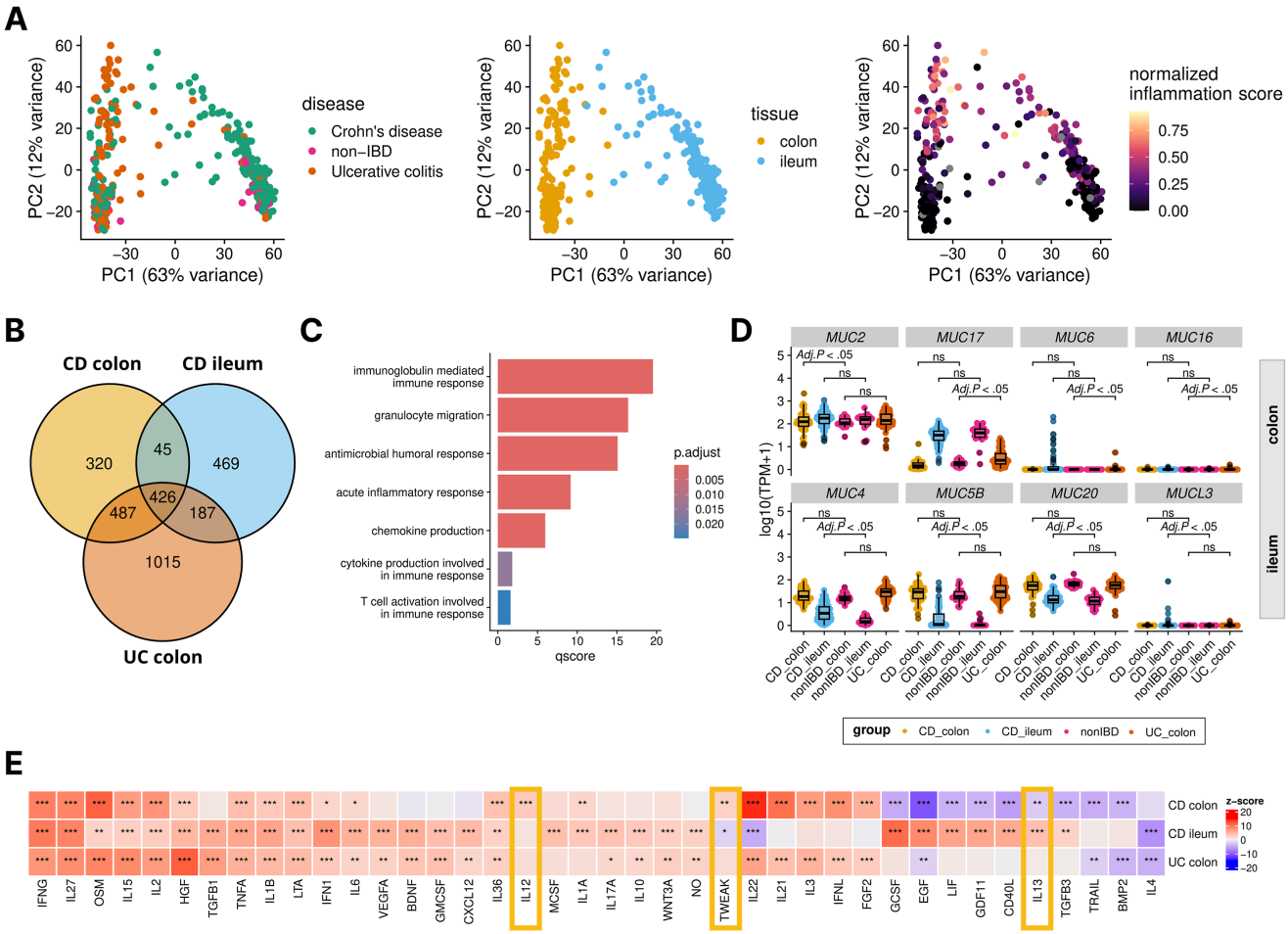


Figure 3. Tissue disease–specific inflammatory gene signatures. (*A*) Principal component (PC) analysis of gene expression data, colored by disease type, tissue, and normalized inflammation as assessed by histopathology (normalized modified Naini Cortina[22] score or normalized modified Riley[21] score). (*B*) *Venn-diagram* depicting the overlap of differentially expressed (DE) genes in the different comparisons (CD inflamed colon vs non-IBD colon; CD inflamed ileum vs non-IBD ileum, and UC inflamed colon vs non-IBD colon). (*C*) Commonly up-regulated Gene Ontology-Biological Process terms across all groups. (*D*) Expression [log10(TPM + 1)] of significantly up-regulated MUCs detected by DE analysis; adjusted *P* values were derived from the DE analysis with DESeq2. (*E*) Cytokine signaling activities in the different groups inferred with CytoSig; *z* scores and *P* values were derived with the CytoSig permutation test. *False discovery rate (FDR) $<$ 0.1; **FDR $<$ 0.05; ***FDR $<$ 0.01.

differential gene expression analyses relative to the corresponding non-IBD control groups (non-IBD colon and non-IBD ileum) (Supplementary Figure 4*A*, Supplementary Tables 10–12). The overlapping differentially expressed genes (adjusted *P* value $<$ .05 and |log2 fold-change| $>$ 1) are shown in Figure 3*B* and Supplementary Figure 4*B*. An overrepresentation analysis of these significantly up-regulated genes, using the Gene Ontology-Biological Process database, revealed enrichment for known immune-related pathways, including acute inflammatory response, chemokine, and cytokine production (Figure 3*C*). Over-representation analysis of overlapping down-regulated genes did not show enrichment for any term, but the expression profiles are shown in Supplementary Figure 4*C*, highlighting the differences in gene expression between tissues.

The composition of the mucus layer varies between the colon and ileum,[39] and previous studies have shown that the structure and function of the mucosal barrier, including the mucus layer, may be significantly disrupted in IBD.[40,41] Mucins (MUCs), which are proteins expressed by epithelial cells, are key components of the mucus. Differential gene expression analysis revealed that 7 MUCs and 1 MUC-like gene were significantly up-regulated (adjusted *P* value $<$ .05 and |log2 fold-change| $>$ 1): *MUC2* in the colon of patients with CD; *MUC6*, *MUC16*, and *MUC17* in the colon of patients with UC; and *MUC5B*, *MUC4*, *MUC20*, and *MUCL3* in the ileum of patients with CD (Figure 3*D*). *MUC6*, *MUC16*, and *MUCL3* are generally expressed at low levels and are therefore likely to be of limited relevance. In contrast, *MUC17*, a transmembrane MUC found in both the colon and small intestine, is significantly up-regulated in inflamed UC colon samples compared with non-IBD controls, but no significant changes were observed in CD. Interestingly, we also observed an up-regulation of *MUC4* in inflamed ileal CD samples, although *MUC4* is primarily associated with colonic membrane MUCs.

To better understand the signaling pathways involved in IBD, we inferred cytokine signaling activities using CytoSig software.[42] The majority (n = 40) of cytokine signaling pathways encoded within CytoSig (total n = 43) were significantly activated or suppressed in at least 1 of the site-specific conditions (Figure 3*E*). The most commonly known pathways, such as TNF-α, OSM, and interferon (IFN) gamma, showed consistently high activation in all inflamed samples compared with non-IBD controls. Notably, we also identified site-specific pathway activations, including IL22, IL21, IL3, IFN lambda, and fibroblast growth factor 2, in inflamed colon samples regardless of disease entity. In addition, we observed disease subtype–specific pathway dysregulation, such as the IL13 pathway in CD (up ileum and down in colon), but not in UC. This aligns with the failure of anti-IL13 antibody therapies in clinical trials in UC.[43,44] Two signaling pathways, IL12 and, to a lesser extent, TNF-like weak inducer of apoptosis, were significantly active in inflamed colonic CD samples (Figure 3*E*). Examining the expression of individual genes involved in IL12 signaling (Supplementary Figure 4*D*), we observed a modest, but statistically significant increase in the expression of *IL12A*, *IL12B*, and *IL12RB2* in inflamed colonic samples from patients with CD compared with colonic non-IBD control samples. Consistent with our findings, Dulai et al[45] reported in a meta-analysis of the CERTIFI and UNITI clinical trials that treatment with ustekinumab, an IL12- and IL-23p40 antibody, was less effective in patients with CD with isolated ileal compared with colonic disease.

Multiomics Profiling Identifies Serum Protein Biomarkers for Disease Localization in Inflammatory Bowel Disease

We next sought to determine whether distinct immunologic pathways driving IBD can be leveraged to identify biomarkers capable of differentiating disease subgroups. Such biomarkers could provide a basis for improved diagnosis, stratification, and personalized treatment strategies for patients with IBD.[9] We categorized serum protein samples into 3 groups based on inflammatory disease localization: CD-ileum (isolated ileal disease), CD-colon, and UC-colon. We then performed a differential protein abundance analysis comparing samples from patients with IBD with active inflammation against non-IBD controls (Figure 4*A*, Supplementary Tables 13–15). This analysis identified 5 proteins—TNF, IL12B, AXIN1, OSM, and TNF superfamily 14—that were commonly up-regulated in all patient groups. Colon samples of both IBD entities showed the highest overlap of differentially abundant proteins (n = 8: CC-chemokine ligand [CCL] 20, CCL25, CXCL1, CXCL11, EN-RAGE, hepatocyte growth factor [HGF], IL24, and latency-associated peptide transforming growth factor [TGF]-β1), although no commonly regulated proteins were identified between ileal CD and colonic UC (Figure 4*B*, Supplementary Figure 5*A*). In ileal CD, the uniquely regulated proteins CUB domain-containing protein 1, leukemia inhibitory factor receptor, and C-X3-C motif chemokine ligand 1 were all down-regulated in patients with active inflammation compared with non-IBD controls.

To explore potential associations between severity of inflammation and protein abundance, we integrated protein data with histopathology inflammatory scores of both IBD entities (modified Naini Cortina[22] score for CD and modified Riley[21] score for UC). In UC, all 6 up-regulated serum proteins, TGF-α, matrix metalloproteinase (MMP) 10, CCL11, IL10, IL17A, and IL7 (Figure 4*B*), showed significant positive correlations with the modified Riley score (Figure 4*C*). Conversely, only 1 protein exhibited a significant correlation with the modified Naini Cortina score in colonic CD (ie, SLAMF1) (Figure 4*C*). Notably, most colon-specific proteins (shared between colonic CD and UC) were positively correlated with the histologic inflammation scores, with the exception of 2 proteins—CCL25 and EN-RAGE (Supplementary Figure 5*A* and *B*). Among the overlapping proteins in CD, an increased abundance of IFN gamma and decreased abundance of fibroblast growth factor 19 and CCL4 were observed. However, only IFN gamma displayed a significant correlation with the severity of inflammation (Figure 4*C*). Mucosal expression of *IFNG* is known to be up-regulated in inflamed CD.[46]

Building on these findings, we next examined the association between protein abundance in the serum and tissue gene expression. Across all samples, the strongest correlation between protein abundance and tissue gene expression was observed for CXCL9 (Pearson's $r = 0.4$) and the strongest inverse correlation for IL2 (Pearson's $r = -0.4$) (Supplementary Figure 5*C*). Stratification of samples by disease and site revealed several significant correlations (Supplementary Figure 5*C*), such as MMP-10, IL17A, and TGF-α (inverse correlation) in UC (Figure 4*D*).

In summary, we identified 6 proteins (ie, CCL20, CXCL1, CXCL11, HGF, IL24, and latency-associated peptide TGF-β1) with increased abundance in colonic diseases, irrespective of the disease entity (colonic CD and UC) that correlated with histologic inflammatory severity. MMP-10, IL17A, and TGF-α were more prominently associated with UC, and elevated serum IFN gamma was linked to CD (Figure 4*E*). Importantly, only the UC-associated markers (ie, MMP-10, IL17A, and TGF-α) also showed significant correlations with tissue gene expression. These findings align with previous research showing higher tissue gene expression levels of *MMP10* in active UC compared with active colonic CD and controls, as well as an association with disease activity in UC.[47] Similarly, multiple studies have reported elevated HGF serum levels and mucosal gene expression in IBD, particularly in UC.[48,49]

Artificial Intelligence Foundation Models Predict Histologic Disease Activities Accurately

In an attempt to develop a robust predictor for histologic disease activity scores, directly from pathology images of intestinal mucosal sections, we applied foundation models on images of H&E-stained tissues (Figure 5*A*) to predict the modified Naini Cortina[22] and modified Riley[21] scores. We used the following 4 different foundation models: CHIEF,[24] UNI2,[25] Virchow2,[26] and H-optimus-1[27] and applied an attention-based multiple instance learning

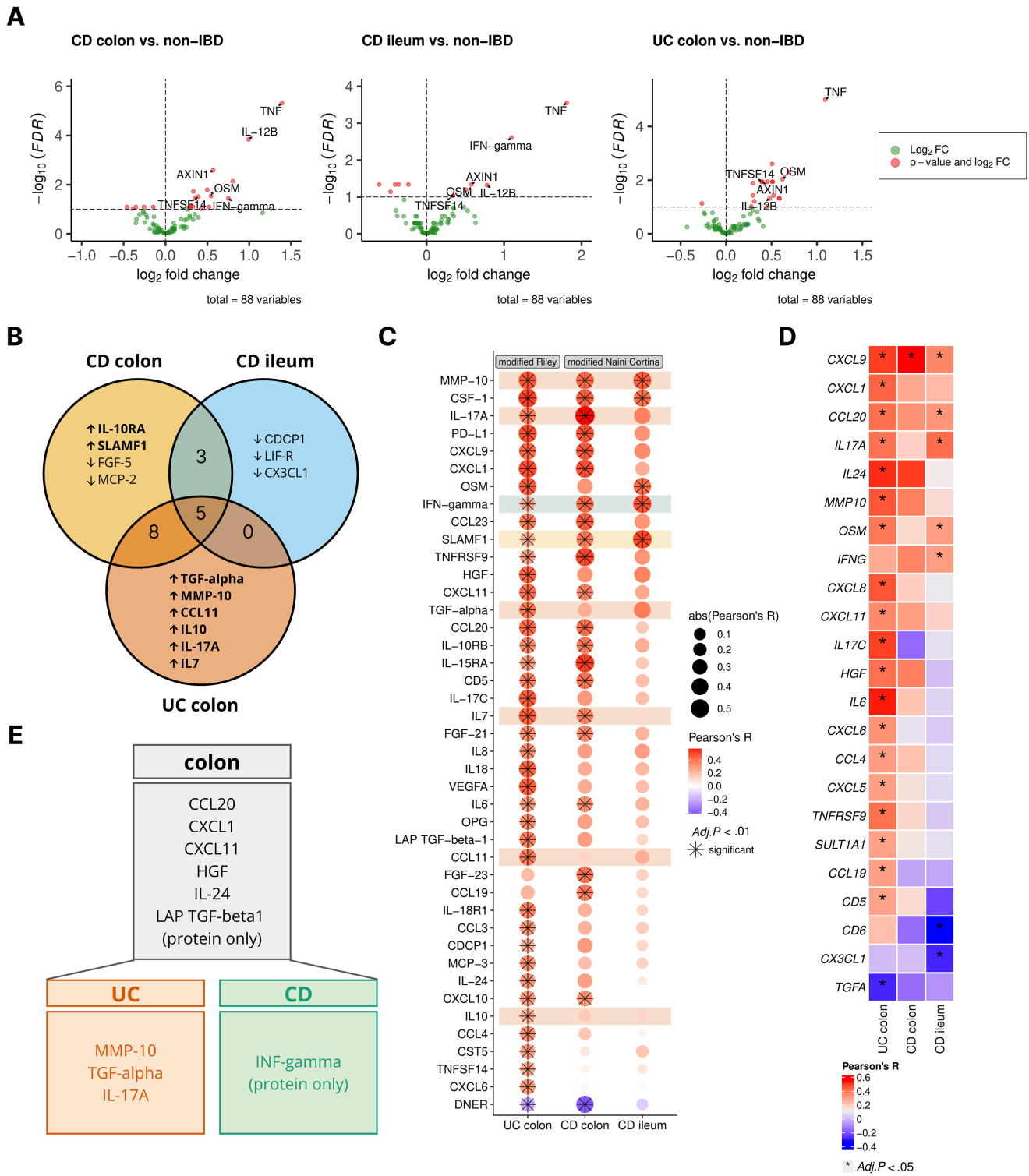


Figure 4. Multiomics profiling identifies potential serum protein biomarkers for disease localization in IBD. (*A*) *Volcano plots* displaying differentially abundant proteins in inflamed, disease site–specific groups compared with non-IBD controls. Statistical significance was determined using Welch's *t* test with Benjamini-Hochberg correction (false discovery rate [FDR] < 0.1). (*B*) *Venn diagram* illustrating the overlap of significantly differentially abundant proteins among CD colon, CD ileum, and UC colon, relative to non-IBD controls. (*C*) *Dot plot* showing Pearson correlation coefficients (*r*) between serum protein abundance and histopathology scores (modified Riley[21] score for UC, modified Naini Cortina[22] score for CD) across the 3 subgroups. Highlighted are uniquely identified differentially abundant proteins from (*A*) and (*B*). Significance threshold: adjusted *P* value < .01. (*D*) *Heatmap* of Pearson correlation coefficients between serum protein abundance and tissue gene expression in the different groups. *Adjusted *P* value < .05. (*E*) Potential serum proteins associated with colonic disease, UC, and CD that significantly correlate with histopathology scores and, with the exception of IFN gamma, also with tissue gene expression.

model to predict histologic disease activity scores (Figure 5*A*). To evaluate the prediction performance, we used 1212 H&E images and categorized them according to histologic disease activity scores: 699 images with the modified Naini Cortina score (514 images from Berlin and 185 from Erlangen) and 556 with the modified Riley score (472 images from Berlin and 84 from Erlangen) (Supplementary Figure 6*A*). We performed a 5-fold cross-validation using the Berlin cohort to train and internally validate the model.

The performance of the different foundation models was assessed based on Pearson correlation between true and predicted scores (Figure 5*B*). The highest performance in predicting the normalized modified Riley[21] score was achieved by the Virchow2 model ($r = 0.933$), while the UNI2 model showed the best results for the modified Naini Cortina[22] score ($r = 0.801$) (Supplementary Figure 6*B*). We then deployed the models to the Erlangen cohort that showed strong performance for both scores (Supplementary Figure 6*C*). We further validated the models on the Innsbruck cohort (Supplementary Figure 7*A*), resulting in high agreement with multiobserver consensus histologic disease activity scores (Supplementary Figure 7*B*). Notably, both the Erlangen and Innsbruck cohorts consisted almost exclusively of biopsy samples, which provide more limited tissue context compared with surgical resection specimens. Moreover, correlations between original and predicted scores with various scoring systems were assessed, with the AI-based scores matching or even surpassing original scores, offering a viable alternative scoring method (Supplementary Figure 6*D*–*F*).

To understand the decision-making process of the AI model, we generated attention heatmaps to highlight the most influential regions for the model's predictions. These heatmaps were then reviewed by expert pathologists. In Figure 5*C*, the heatmap of a patient with UC shows the model's attention levels. Regions with high attention indicate strong influence on the model's prediction, focusing primarily on peripheral areas near the mucosa and submucosa lining. These regions often display histologic signs of disease activity, such as crypt abscesses, immune cell infiltration, architectural distortion, and signs of increased epithelial regeneration—hallmarks of UC pathology. This is demonstrated by 4 of the top attention tiles (Figure 5*D*), which highlight areas with inflammatory cell infiltration, including lymphocytes and plasma cells as well as distorted crypts and crypt abscesses. In contrast, low-attention regions are concentrated in the inner, noninflamed, mucosal areas. Importantly, the model did not consider components of the immune environment such as lymph follicles and lymph nodes. Supplementary Figure 6*G* provides an additional example from a patient with CD with moderate disease activity, where the model similarly focuses on pathologically relevant regions.

In summary, foundation models enable a robust prediction and assessment of histologic disease activity scores. Moreover, the models accurately identified histologic patterns consistent with UC pathology when predicting disease activity.

Foundation Models Enable Accurate Classification of Inflammatory Bowel Disease Subtypes

To evaluate whether foundation models-derived embeddings extracted from H&E slides can capture IBD subtype–relevant histopathologic information, we benchmarked the same model in a binary classification task using both Berlin and Erlangen cohorts (Figure 6*A*). We trained the models on the full Berlin dataset (n = 970) and deployed them on the Erlangen cohort, including both inflamed and noninflamed samples (Figure 6*A*). Validation was carried out on multiple deployment subsets: all inflamed tissues (n = 132), colon-only samples (n = 125), and inflamed colon samples (n = 52) (Supplementary Figure 8*A*). Virchow2 achieved the best performance, particularly on the inflamed tissue subset of the Erlangen cohort, reaching an area under the receiver operating characteristic of 0.93 (95% CI, 0.88–0.97) and an area under the precision-recall curve of 0.97 (95% CI, 0.94–0.99) (Figure 6*B*). Similar performance was also achieved on the full Erlangen cohort (area under the receiver operating characteristic, 0.87; 95% CI, 0.83–0.91 and area under the precision-recall curve, 0.95; 95% CI, 0.93–0.97) (Supplementary Figure 8*B*). Confusion matrices summarizing classification accuracy in both settings are shown in Figure 6*C* (inflamed only) and Supplementary Figure 8*C* (full cohort). These findings demonstrate that foundation model features enable robust and generalizable discrimination of IBD subtypes across different patient cohorts.

Sample numbers for each subgroup in all analyses are summarized in Supplementary Table 16.

Discussion

We created a comprehensive molecular, histopathologic, and clinical atlas of IBD by profiling more than 1000 patients using multiomic and multimodal assays. Generation and integration of genomic, transcriptomic, serum proteomic, and H&E histologic imaging data, coupled with standardized clinical disease characteristics annotation data, including histopathology and endoscopy scores, make IBDome a comprehensive resource for IBD. To enhance the exploitation of this resource, we are providing a publicly available web platform for data exploration, analysis, and validation (https://ibdome.org). Beyond building this freely accessible resource for scientific research, our study provides several important insights.

First, we developed an inflammatory protein signature from serum samples that reflects the underlying intestinal inflammation and can be used to monitor disease activity of patients noninvasively. The IBD-IPSS provides a novel approach to assess disease severity, complementing existing molecular and clinical scores. Our findings demonstrate that this serum-based signature strongly correlates with established endoscopic scores, underscoring its potential as a biomarker for disease monitoring. The identification of OSM as the only overlapping protein between the IBD-IPSS and the circulating molecular inflammation score[31]

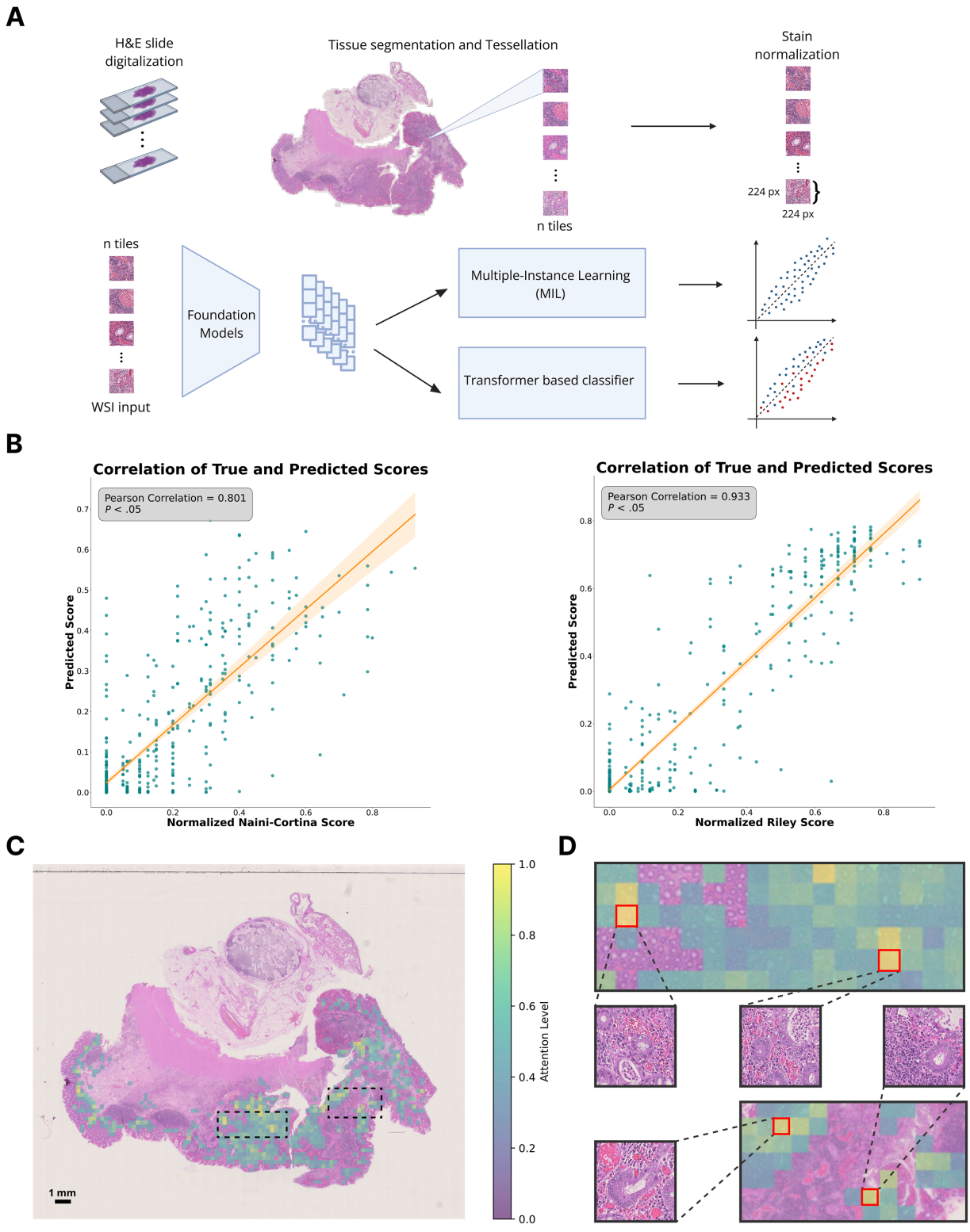


Figure 5. Prediction of histologic disease activity from pathology images. (*A*) Overview of the AI-based histopathology analysis pipeline. (*B*) Correlation plots between the original histologic disease activity scores (*x-axis*) and artificial intelligence–predicted scores (*y-axis*) for both scoring systems, based on 5-fold cross-validation on the Berlin subset using the best performing Foundation Model (UNI2 and Virchow2, respectively). (*C*) Representative attention *heatmap* of a WSI from a patient with UC with high histologic disease activity. The *heatmap* shows the model's attention levels, displaying only tiles with scores $>$ 0.4. Higher scores (*yellow*) mark regions that strongly influence the model's prediction, while lower scores (*green*) indicate less critical regions. (*D*) Zoomed-in view of the highest-attention regions highlighted in (*C*), showing 4 of the top 10 attention tiles, outlined in *red*.

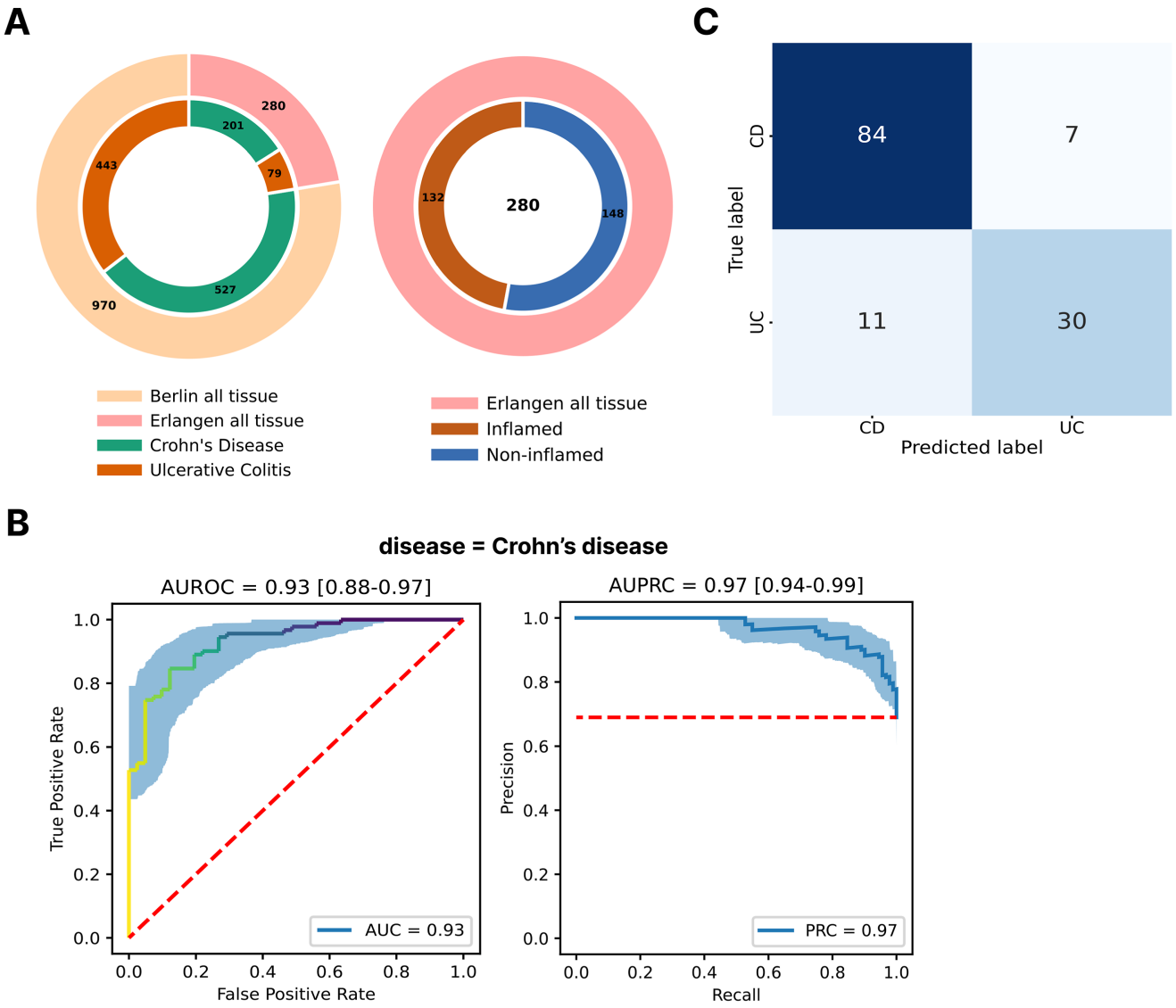


Figure 6. Prediction of disease type from pathology images. (*A*) Cohort distribution: on the *left*, the *outer circle* represents the distribution across centers and the *inner circle* shows the proportions of CD and UC cases. On the *right*, the *outer circle* shows the Erlangen all-tissue subset, and the *inner circle* indicates the proportion of inflamed and noninflamed samples. (*B*) Area under the receiver operating characteristic curve (AUROC) and area under the precision-recall curve (AUPRC) for the best-performing model (Virchow2) trained on the Berlin cohort and evaluated on the Erlangen inflamed all-tissue subset. (*C*) Confusion matrix summarizing the UC/CD classification results from the deployment in (*B*).

suggests its central role in systemic inflammation and further supports its relevance in IBD pathophysiology.[32]

Second, we uncovered distinct site-specific inflammatory signatures of CD and UC, emphasizing that the disease site plays a crucial role in shaping the inflammatory landscape. The differential gene expression of MUCs provides further insight into the tissue-specificity of IBD pathology. The selective up-regulation of *MUC17* in UC colon inflammation but not in CD, and the increased expression of *MUC4* in inflamed CD ileum, suggest distinct mechanisms of barrier dysfunction in different disease subtypes. Moreover, our cytokine signaling analysis revealed key differences in inflammatory pathway activation across disease subtypes and sites. Although canonical inflammatory pathways such as TNF-α and OSM were consistently up-regulated in all inflamed tissues, we identified site-specific and disease subtype–specific pathway activations, including IL12 signaling in colonic CD. This is particularly relevant, given the variable response to biologic therapies targeting IL12/23, such as ustekinumab, which has been found to be less effective in isolated ileal CD compared with colonic CD.[45] However, the magnitude of benefit with anti-IL23p19 antagonists is also numerically higher in patients with isolated colonic vs ileal disease, indicating that target expression is not the sole reasoning underlying heightened therapeutic effectiveness.[50]

At the serum protein level, we observed that colonic CD and UC share a substantial overlap in differentially abundant proteins, while ileal CD exhibits a more distinct inflammatory profile. The ability to differentiate IBD subtypes based on serum protein signatures offers a promising avenue for noninvasive disease monitoring and personalized treatment approaches. Specifically, the

detection of MMP-10, IL17A, and TGF-α as UC-associated markers and IFN gamma as a CD-associated marker may help in more accurate disease classification and targeted therapeutic strategies. Given the failure of anti-IL13 therapies in CD[43,44] and the ongoing investigation of anti–IFN-gamma antibodies,[51,52] our results emphasized the need to guide treatment strategies based on disease localization and immune signatures.

Third, we found that foundation models for images of H&E-stained tissue sections exhibit high predictive performance in approximating expert-assigned histologic disease activity scores and classifying disease subtype. Our deep learning framework demonstrated high correlation between predicted and true scores. Moreover, explainability analyses showed that the model focuses on histologically relevant regions aligning closely with expert pathologist assessments. Furthermore, the model's predictions showed a strong correlation with endoscopic scoring systems, as well as molecular scores. These findings suggest that AI-based histologic scoring could help to reduce interobserver variability by supporting pathologists in achieving a more standardized and objective assessment of disease activity, thereby improving objective disease monitoring in IBD.

Our work has some limitations that should be acknowledged. Although the multiomic profiling enabled broad molecular characterization, serum proteomic analyses were restricted to the 96-plex Olink panel, and the cross-sectional study design limits conclusions on temporal dynamics and outcome prediction. In addition, the non-IBD control group was moderately older than the patients with IBD.

In conclusion, the IBDome is a powerful resource for uncovering IBD biology and ultimately advancing precision medicine to improve patient outcomes.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at http://doi.org/10.1053/j.gastro.2026.05.023.

References

1. Le Berre C, Ananthakrishnan AN, Danese S, et al. Ulcerative colitis and Crohn's Disease have similar burden and goals for treatment. Clin Gastroenterol Hepatol 2020;18:14–23.
2. Sato Y, Tsujinaka S, Miura T, et al. Inflammatory bowel disease and colorectal cancer: epidemiology, etiology, surveillance, and management. Cancers 2023;15:4154.
3. Wilson JC, Furlano RI, Jick SS, et al. Inflammatory bowel disease and the risk of autoimmune diseases. J Crohns Colitis 2016;10:186–193.
4. Wang R, Li Z, Liu S, et al. Global, regional and national burden of inflammatory bowel disease in 204 countries and territories from 1990 to 2019: a systematic analysis based on the Global Burden of Disease Study 2019. BMJ Open 2023;13:e065186.
5. Dolinger M, Torres J, Vermeire S. Crohn's disease. Lancet 2024;403:1177–1191.
6. Scott FI, Ananthakrishnan AN, Click B, et al. AGA living clinical practice guideline on the pharmacologic management of moderate-to-severe Crohn's disease. Gastroenterology 2025;169:1397–1448.
7. Singh S, Loftus EV, Limketkai BN, et al. AGA living clinical practice guideline on pharmacological management of moderate-to-severe ulcerative colitis. Gastroenterology 2024;167:1307–1343.
8. Turner D, Ricciuto A, Lewis A, et al. STRIDE-II: an update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) Initiative of the International Organization for the Study of IBD (IOIBD): determining therapeutic goals for treat-to-target strategies in IBD. Gastroenterology 2021;160:1570–1583.
9. Atreya R, Neurath MF. Biomarkers for personalizing IBD therapy: the quest continues. Clin Gastroenterol Hepatol 2024;22:1353–1364.
10. Gordon H, Minozzi S, Kopylov U, et al. ECCO guidelines on therapeutics in Crohn's disease: medical treatment. J Crohns Colitis 2024;18:1531–1555.
11. Proctor LM, Creasy HH, Fettweis JM, et al. The Integrative Human Microbiome Project. Nature 2019;569:641–648.
12. Lloyd-Price J, Arze C, Ananthakrishnan AN, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature 2019;569:655–662.
13. Imhann F, Van der Velde KJ, Barbieri R, et al. The 1000IBD project: multi-omics data of 1000 inflammatory bowel disease patients; data release 1. BMC Gastroenterol 2019;19:5.
14. **Massimino L, Lamparelli LA**, Houshyar Y, et al. The Inflammatory Bowel Disease Transcriptome and Metatranscriptome Meta-Analysis (IBD TaMMA) framework. Nat Comput Sci 2021;1:511–515.
15. Kong L, Pokatayev V, Lefkovith A, et al. The landscape of immune dysregulation in Crohn's disease revealed through single-cell transcriptomic profiling in the ileum and colon. Immunity 2023;56:444–458.e5.
16. Smillie CS, Biton M, Ordovas-Montanes J, et al. Intra- and inter-cellular rewiring of the human colon during ulcerative colitis. Cell 2019;178:714–730.e22.
17. **D'Haens G, Sandborn WJ**, Feagan BG, et al. A review of activity indices and efficacy end points for clinical trials of medical therapy in adults with ulcerative colitis. Gastroenterology 2007;132:763–786.
18. Harvey RF, Bradshaw JM. A simple index of Crohn's-disease activity. Lancet 1980;1:514.
19. Xie T, Zhang T, Ding C, et al. Ulcerative Colitis Endoscopic Index of Severity (UCEIS) versus Mayo Endoscopic Score (MES) in guiding the need for colectomy in patients with acute severe colitis. Gastroenterol Rep 2018;6:38–44.
20. Daperno M, D'Haens G, Van Assche G, et al. Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD. Gastrointest Endosc 2004;60:505–512.
21. Riley SA, Mani V, Goodman MJ, et al. Microscopic activity in ulcerative colitis: what does it mean? Gut 1991;32:174–178.

22. Naini BV, Cortina G. A histopathologic scoring system as a tool for standardized reporting of chronic (ileo)colitis and independent risk assessment for inflammatory bowel disease. Hum Pathol 2012;43:2187–2196.
23. Macenko M, Niethammer M, Marron JS, et al. A method for normalizing histology slides for quantitative analysis. In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, 2009:1107–1110.
24. **Wang X, Zhao J**, Marostica E, et al. A pathology foundation model for cancer diagnosis and prognosis prediction. Nature 2024;634:970–978.
25. **Chen RJ, Ding T, Lu MY**, et al. Towards a general-purpose foundation model for computational pathology. Nat Med 2024;30:850–862.
26. **Vorontsov E, Bozkurt A, Casson A**, et al. A foundation model for clinical-grade computational pathology and rare cancers detection. Nat Med 2024;30:2924–2935.
27. Scalbert M, Saillard C, Peeters T, et al. Abstract LB174: H-optimus-1: a foundation model for computational histopathology. Cancer Res 2026;86:LB174–LB174.
28. Leiby JS, Hao J, Kang GH, et al. Attention-based multiple instance learning with self-supervision to predict microsatellite instability in colorectal cancer from histology whole-slide images. In: *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2022:3068–3071.
29. El Hadad J, Schreiner P, Vavricka SR, et al. The genetics of inflammatory bowel disease. Mol Diagn Ther 2024;28:27–35.
30. Finotello F, Mayer C, Plattner C, et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. Genome Med 2019;11:34.
31. Argmann C, Hou R, Ungaro RC, et al. Biopsy and blood-based molecular biomarker of inflammation in IBD. Gut 2023;72:1271–1287.
32. West NR, Hegazy AN, Owens BMJ, et al. Oncostatin M drives intestinal inflammation and predicts response to tumor necrosis factor-neutralizing therapy in patients with inflammatory bowel disease. Nat Med 2017;23:579–589.
33. Yau TO, Vadakekolathu J, Foulds GA, et al. Hyperactive neutrophil chemotaxis contributes to anti-tumor necrosis factor-α treatment resistance in inflammatory bowel disease. J Gastroenterol Hepatol 2022;37:531–541.
34. Mudter J, Neurath MF. IL-6 signaling in inflammatory bowel disease: pathophysiological role and clinical relevance. Inflamm Bowel Dis 2007;13:1016–1023.
35. Belarif L, Danger R, Kermarrec L, et al. IL-7 receptor influences anti-TNF responsiveness and T cell gut homing in inflammatory bowel disease. J Clin Invest 2019;129:1910–1925.
36. Williams MA, O'Callaghan A, Corr SC. IL-33 and IL-18 in inflammatory bowel disease etiology and microbial interactions. Front Immunol 2019;10:1091.
37. **Atreya R, Siegmund B**. Location is important: differentiation between ileal and colonic Crohn's disease. Nat Rev Gastroenterol Hepatol 2021;18:544–558.
38. Cleynen I, Boucher G, Jostins L, et al. Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. Lancet 2016;387:156–167.
39. Johansson MEV, Hansson GC. Immunological aspects of intestinal mucus and mucins. Nat Rev Immunol 2016;16:639–649.
40. Buisine M-P, Desreumaux P, Debailleul V, et al. Abnormalities in mucin gene expression in Crohn's disease. Inflamm Bowel Dis 1999;5:24–32.
41. Leoncini G, Cari L, Ronchetti S, et al. Mucin expression profiles in ulcerative colitis: new insights on the histological mucosal healing. Int J Mol Sci 2024;25:1858.
42. **Jiang P, Zhang Y, Ru B**, et al. Systematic investigation of cytokine signaling activity at the tissue and single-cell levels. Nat Methods 2021;18:1181–1191.
43. Danese S, Rudziński J, Brandt W, et al. Tralokinumab for moderate-to-severe UC: a randomised, double-blind, placebo-controlled, phase IIa study. Gut 2015;64:243–249.
44. Reinisch W, Panés J, Khurana S, et al. Anrukinzumab, an anti-interleukin 13 monoclonal antibody, in active UC: efficacy and safety from a phase IIa randomised multicentre study. Gut 2015;64:894–900.
45. Dulai PS, Singh S, Vande Casteele N, et al. Should we divide Crohn's disease into ileum-dominant and isolated colonic diseases? Clin Gastroenterol Hepatol 2019;17:2634–2643.
46. Neurath MF. Cytokines in inflammatory bowel disease. Nat Rev Immunol 2014;14:329–342.
47. **Fonseca-Camarillo G, Furuzawa-Carballeda J**, Martínez-Benitez B, et al. Increased expression of extracellular matrix metalloproteinase inducer (EMMPRIN) and MMP10, MMP23 in inflammatory bowel disease: cross-sectional study. Scand J Immunol 2021;93:e12962.
48. Naguib R, El-Shikh WM. Clinical significance of hepatocyte growth factor and transforming growth factor-beta-1 levels in assessing disease activity in inflammatory bowel disease. Can J Gastroenterol Hepatol 2020;2020:2104314.
49. Stakenborg M, Verstockt B, Meroni E, et al. Neutrophilic HGF-MET signalling exacerbates intestinal inflammation. J Crohns Colitis 2020;14:1748–1758.
50. Lee HH, Yuan Y, Boland BS, et al. Efficacy of advanced therapies in achieving remission by disease location in Crohn's disease: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2026;24:923–931.
51. Hommes DW, Mikhajlova TL, Stoinov S, et al. Fontolizumab, a humanised anti-interferon γ antibody, demonstrates safety and clinical activity in patients with moderate to severe Crohn's disease. Gut 2006;55:1131–1137.
52. Reinisch W, de Villiers W, Bene L, et al. Fontolizumab in moderate to severe Crohn's disease: a phase 2, randomized, double-blind, placebo-controlled, multiple-dose study. Inflamm Bowel Dis 2010;16:233–242.

Author names in bold designate shared co-first authorship.

Received September 12, 2025. Accepted May 20, 2026.

Correspondence
Address correspondence to: Christoph Becker, PhD, Department of Medicine 1, Friedrich-Alexander-Universität Erlangen-Nürnberg and Universitätsklinikum Erlangen, Erlangen, Germany. e-mail: christoph.becker@uk-erlangen.de; or Britta Siegmund, MD, Medizinische Klinik für Gastroenterologie, Infektiologie

und Rheumatologie, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany. e-mail: britta.siegmund@charite.de; or Zlatko Trajanoski, PhD, Biocenter, Institute of Bioinformatics, Medical University of Innsbruck, Innsbruck, Austria. e-mail: zlatko.trajanoski@i-med.ac.at.

Acknowledgments

The TRR241 IBDome Consortium includes (in alphabetical order) Imke Atreya,[1,2] Raja Atreya,[1,2] Petra Bacher,[3,4] Christoph Becker,[1,2] Christian Bojarski,[5] Nathalie Britzen-Laurent,[6] Caroline Voskens,[7] Hyun-Dong Chang,[8,9] Andreas Diefenbach,[10] Claudia Günther,[1,2] Ahmed N. Hegazy,[5] Kai Hildner,[1,2] Christoph S. N. Klose,[10] Kristina Koop,[1,2] Susanne M. Krug,[11] Anja A. Kühl,[12] Moritz Leppkes,[1,2] Rocío López-Posadas,[1,2] Leif S. Ludwig,[13,14] Clemens Neufert,[1,2] Markus Neurath,[1,2] Jay V. Patankar,[1,2] Christina Plattner,[15] Magdalena S. Prüß,[5] Andreas Radbruch,[8] Chiara Romagnani,[16] Francesca Ronchi,[10] Ashley D. Sanders,[5,13,14] Alexander Scheffold,[4] Jörg-Dieter Schulzke,[11] Michael Schumann,[5] Sebastian Schürmann,[17] Britta Siegmund,[5] Michael Stürzl,[6] Zlatko Trajanoski,[15] Antigoni Triantafyllopoulou,[8,10,18] Maximilian Waldner,[1,2] Carl Weidinger,[5] Stefan Wirtz,[1,2] and Sebastian Zundler[1,2]; from the [1]Department of Medicine 1, Universitätsklinikum, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany; [2]Deutsches Zentrum Immuntherapie, Universitätsklinikum, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany; [3]Institute of Clinical Molecular Biology, Christian-Albrecht University of Kiel, Kiel, Germany; [4]Institute of Immunology, Christian-Albrecht University of Kiel and University Hospital Schleswig-Holstein, Kiel, Germany; [5]Department of Gastroenterology, Infectious Diseases and Rheumatology, Charité–Universitätsmedizin Berlin, Berlin, Germany; [6]Department of Surgery, Universitätsklinikum, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany; [7]Department of Dermatology, Universitätsklinikum, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany; [8]Deutsches Rheuma-Forschungszentrum, ein Institut der Leibniz-Gemeinschaft, Berlin, Germany; [9]Institute of Biotechnology, Technische Universität Berlin, Germany; [10]Institute of Microbiology, Infectious Diseases and Immunology, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; [11]Clinical Physiology and Nutritional Medicine, Department of Gastroenterology, Rheumatology and Infectious Diseases, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; [12]iPATH.Berlin, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; [13]Berlin Institute of Health at Charité–Universitätsmedizin Berlin, Berlin, Germany; [14]Max Delbrück Center für Molekulare Medizin, Berlin, Germany; [15]Biocenter, Institute of Bioinformatics, Medical University of Innsbruck, Innsbruck, Austria; [16]Institute for Medical Immunology, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; [17]Institute of Medical Biotechnology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany; and [18]Department of Rheumatology and Clinical Immunology, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany.

The authors would like to thank Matthias Ziehm and Christin Beier from the Core Unit Proteomics, Berlin Institute of Health at Charité, Berlin for the Olink proteomics measurements, quality control, and technical advice; Sebastian Schönherr from the Institute of Genetic Epidemiology, Medical University of Innsbruck, for his advice on performing genotype imputation, and Ludmilla Sologub, Angelika Wilfer and Isabella Dörflein for excellent technical assistance. The graphical abstract was created with BioRender.com.

CrediT Authorship Contributions

Christina Plattner, PhD (Formal analysis: Lead; Software: Lead; Writing – original draft: Equal; Writing – review & editing: Equal)

Gregor Sturm, PhD (Formal analysis: Lead; Software: Lead; Writing – original draft: Equal; Writing – review & editing: Equal)

Anja A. Kühl, PhD (Formal analysis: Equal; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Raja Atreya, MD (Formal analysis: Equal; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Sandro Carollo, MS (Formal analysis: Equal; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Dietmar Rieder, PhD (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Raphael Gronauer, PhD (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Michael Günther, MD (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Steffen Ormanns, MD (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Claudia Manzl, PhD (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Stefan Wirtz, PhD (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Asier Rabasco Meneghetti, PhD (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Ahmed N. Hegazy, MD (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Jay V. Patankar, PhD (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Zunamys I. Carrero, PhD (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Felix Grabherr, MD (Investigation: Supporting)

Moritz Meyer, MD (Investigation: Supporting)

Timon E. Adolph, MD, PhD (Investigation: Supporting)

Herbert Tilg, MD (Investigation: Supporting)

Markus F. Neurath, MD (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Jakob Nikolas Kather, MD (Formal analysis: Supporting; Writing – original draft: Supporting; Writing – review & editing: Supporting)

Christoph Becker, PhD (Conceptualization: Equal; Formal analysis: Equal; Funding acquisition: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Britta Siegmund, MD (Conceptualization: Equal; Formal analysis: Equal; Funding acquisition: Equal; Writing – original draft: Equal; Writing – review & editing: Equal)

Zlatko Trajanoski, PhD (Conceptualization: Equal; Formal analysis: Equal; Funding acquisition: Equal; Supervision: Equal; Writing – original draft: Lead; Writing – review & editing: Lead)

Conflicts of interest

These authors disclose the following: Raja Atreya has served as a speaker, or consultant, or received research grants from AbbVie, Abivax, AlfaSigma Global, AstraZeneca, Bristol Myers Squibb, CED Service GmbH, Celltrion Healthcare, Dr. Falk Pharma, Galapagos NV, Johnson & Johnson, Eli Lilly and Company, Materia Prima, MSD, Pfizer, and Takeda Pharmaceuticals. Jakob Nikolas Kather declares consulting services for Panakeia, AstraZeneca, MultiplexDx, Mindpeak, Owkin, DoMore Diagnostics, and Bioptimus. He holds shares in StratifAI, Synagen AI, Tremont AI, and Ignition Labs; has received an institutional research grant from GSK; and has received honoraria from AstraZeneca, Bayer, Daiichi Sankyo, Eisai Co, Ltd, Janssen, Merck, MSD, Bristol Myers Squibb, Roche, Pfizer, and Fresenius. Britta Siegmund consulted for AbbVie, Abivax, AltruBio, Boehringer Ingelheim, Bristol Myers Squibb, Dr. Falk Pharma, Eli Lilly and Company, Endpoint Health, Falk, Galapagos NV, Gilead, Janssen, Landos Biopharma, Lilly, Materia Prima, PredictImmune, Pfizer, and Takeda Pharmaceuticals; received speaker fees from AbbVie, AlfaSigma, Bristol Myers Squibb, CED Service GmbH, Dr. Falk Pharma, Eli Lilly, MSD, Ferring, Galapagos NV, Janssen, Pfizer, and Takeda Pharmaceuticals; and received grant support from Pfizer (all the money went to an institutional account at Charité). The remaining authors disclose no conflicts.

Funding

This work was funded by the Deutsche Forschungsgemeinschaft (DFG [German Research Foundation]) grant TRR241 375876048 (Z03; Z02; B01; A03 to Zlatko Trajanoski, Anja A. Kühl, Raja Atreya, Britta Siegmund, and Christoph Becker), and DFG research grant 418055832 (Christoph Becker and Britta Siegmund). This research was funded in part by the Austrian Science Fund [10.55776/I3978]. For open access purposes, the author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission. Anja A. Kühl was further supported by SFB1340 372486779 (TP B06). Britta Siegmund is further supported by the DFG projects CRU 5023 (project number 50474582), CRC 1449-B04 and Z02 (project number 431232613); CRC 1340-B06 (project number 372486779); and project number: 418055832. Britta Siegmund and Ahmed N. Hegazy are further supported by the DFG under Germany's Excellence Strategy (EXC 3118/1) project number 533770413. Christoph Becker was further supported by the DFG projects CRU 5024-A03 (project 505539112), CRC 1755-TP14 (project 550296805 and DFG research grants 536563567 and 510624836. Timon E. Adolph is further supported by the European Research Council (MEGI CD 101039320) and the Austrian Science Fund (P33070, FG15, and the MYCOS program of the Medical University of Innsbruck). Jakob Nikolas Kather is supported by the German Cancer Aid DKH (DECADE, 70115166), the German Federal Ministry of Research, Technology and Space BMFTR (PEARL, 01KD2104C; CAMINO, 01EO2101; TRANSFORM LIVER, 031L0312A; TANGERINE, 01KT2302 through ERA-NET Transcan; Come2Data, 16DKZ2044A; DEEP-HCC, 031L0315A; DECIPHER-M, 01KD2420A; NextBIG, 01ZU2402A), the DFG (CRC/TR 412, 535081457; SFB 1709/1 2025, 533056198), the German Academic Exchange Service DAAD (SECAI, 57616814), the German Federal Joint Committee G-BA (TransplantKI, 01VSF21048), the European Union's Horizon Europe research and innovation program (ODELIA, 101057091; GENIAL, 101096312), the European Research Council (NADIR, 101114631), the National Institutes of Health (EPICO, R01 CA263318), and the National Institute for Health and Care Research (NIHR) (Leeds Biomedical Research Centre, NIHR203331). The views expressed are those of the authors and not necessarily those of

the National Health Service, the NIHR, or the Department of Health and Social Care. This work was funded by the European Union. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Data Availability

The data can be interactively explored using the IBDome Explorer (https://ibdome.org), where also the full SQLite research database and individual data tables are available for download. Raw data and complete mutation tables are not made available due to privacy concerns, but inflammatory bowel disease–relevant single-nucleotide variants are included in the IBDome database. The database is deposited in Zenodo (https://doi.org/10.5281/zenodo.18153919). Whole-slide images of the H&E-stained tissue sections are available from the BioImage Archive under accession number S-BIAD1753 (https://doi.org/10.6019/S-BIAD1753). The code for reproducing the results of this study is available on GitHub: https://github.com/orgs/ibdome/repositories.

Gastroenterology, Volume *171*

Supplemental information

IBDome: An Integrated Molecular, Histopathological, and Clinical Atlas of Inflammatory Bowel Diseases

Christina Plattner, Gregor Sturm, Anja A. Kühl, Raja Atreya, Sandro Carollo Dietmar Rieder, Raphael Gronauer, Michael Günther, Steffen Ormanns, Claudia Manzl, Stefan Wirtz, Asier Rabasco Meneghetti, Ahmed N. Hegazy, Jay V. Patankar, Zunamys I. Carrero, Felix Grabherr, Moritz Meyer, TRR241 IBDome ConsortiumTimon E. Adolph, Herbert Tilg, Markus F. Neurath, Jakob Nikolas Kather**Christoph Becker, Britta Siegmund, and Zlatko Trajanoski**

Supplemental material

IBDome: An integrated molecular, histopathological, and clinical atlas of inflammatory bowel diseases

Christina Plattner, Gregor Sturm, Anja A. Kühl, Raja Atreya, Sandro Carollo, Dietmar Rieder, Raphael Gronauer, Michael Günther, Steffen Ormanns, Claudia Manzl, Stefan Wirtz, Asier Rabasco Meneghetti, Ahmed N. Hegazy, Jay V. Patankar, Zunamys I. Carrero, Felix Grabherr, Moritz Meyer, TRR241 IBDome Consortium, Timon E. Adolph, Herbert Tilg, Markus F. Neurath, Jakob Nikolas Kather, Christoph Becker, Britta Siegmund, and Zlatko Trajanoski

Supplementary Methods

Ethics approval and consent to participate

The IBDome was approved by the institutional ethics boards in both Erlangen and Berlin (project identifiers 332-17B and EA1/200/17, respectively). The IBDome is granted permission to collect and share patient samples, clinical and molecular data. All included participants are 18 years or older and have provided informed consent before inclusion into the study.

The non-IBD control subjects consisted of patients who underwent surveillance colonoscopy and did not have signs of active intestinal inflammation upon the performed colonoscopy procedure, and also did not have any history of diagnosed chronic intestinal inflammation. If the samples of the non-IBD control group were taken from surgical gut specimens, it was ensured that they were taken from patients who had to undergo surgery due to the diagnosis of colorectal neoplasia and that the selected site consisted of non-neoplastic and non-inflammatory tissue.

H&E-stained tissue sections from patients in Innsbruck were retrieved from the existing biobank which was approved by the institutional ethics board with the identifiers AN4994 323/4.4 (3806a).

Data management

We distinguish between clinical databases at the study center and a centralized research database. The former was implemented by the IT departments of the study centers in accordance with data protection laws, while the latter only contains non-identifiable information that may be shared publicly according to the ethics approval. In regular intervals, data are transferred from the clinical centers to the central research database located in Innsbruck (Biocenter, Institute of Bioinformatics at the Medical University of Innsbruck).

Study participants were assigned a randomly generated pseudonym when entering the study, which was used to label specimens and samples in the research database. The data related to biomaterials are stored in pseudonymized form in the Starlims biobank management software. Access to the systems (clinical databases and Starlims) was restricted and regulated by an authorization concept.

To ensure data security, all systems are hosted in a secured environment of the university hospital IT infra-structure of Erlangen and Berlin with an information security management

system (ISMS) based on guidelines from the German Federal Office for Information Security. The ISMS specifies procedures and rules within the hospital to define, manage, control, maintain, and continuously improve data security. The documented standard operating procedures for data security and data safety were followed and were checked on a regular basis. The data management fulfills all requirements of the EU General Data Protection Regulation and good scientific practice.

Biomaterial collection, processing and storage

The following specimen are collected from patients in the study:

- Whole blood, collected in heparinized tubes (Vacuette® Greiner Bio-One plasma tube with heparin, Thermo Fisher Scientific) for peripheral blood mononuclear cell isolation as well as K3EDTA tubes (Vacuette® Greiner Bio-One, Thermo Fisher Scientific) for DNA isolation.
- Serum, collected in (Vacuette® Greiner Bio-One Z Serum Sep Clot Activator tubes, Thermo Fisher Scientific).
- Mucosal biopsies collected during endoscopy or after surgery from surgical specimen, stored in test tubes containing RNA protect reagent (RNAprotect Tissue Reagent, Qiagen) for RNA isolation and neutral buffered, 10 % formalin solution (Sigma-Aldrich) for histopathology.
- Surgical resections, including ileocecal resection, hemicolectomy, colectomy, and normal tissue during cancer surgery, where we collected the unaffected tissue at the resection margin for IBDome.
- Stool samples, by providing patients with a stool sample tube containing RNA protect reagent (RNAprotect Tissue Reagent, Qiagen) and a questionnaire to sample stool 3-5 days after endoscopy or surgery.

In brief, samples were processed as follows. Peripheral blood mononuclear leukocytes (PBMC) are isolated from whole blood employing the SepMate™-50 (IVD) tube for density gradient centrifugation (StemCell Technologies). PBMCs are stimulated with PMA/Ionomycin and LPS or left unstimulated for 4 hours. Naïve PBMC (directly after isolation), stimulated PBMC and unstimulated PBMC (with or without brefeldin A) are fixed in Proteomic Stabilizer PROT1 (SMART TUBE Inc.) and stored at -80°C for CyTOF analysis. The supernatants of LPS-stimulated PBMC are stored at -80°C for cytokine analysis. Whole blood from EDTA tubes is stored in 1 mL aliquots at -80°C for DNA isolation. Serum is stored in 1 mL aliquots at -80°C for proteomics (Olink). After incubation of biopsies in RNA protect reagent (RNAprotect Tissue Reagent, Qiagen) overnight at 4°C, biopsies are stored individually at

-80°C until RNA isolation. Formalin-fixed biopsies or resected tissue is processed by and stored at iPATH.Berlin, the core unit of Charité-Universitätsmedizin Berlin for histopathology. Stool samples in RNA protect reagent (RNAprotect Tissue Reagent, Qiagen) are stored in pea-sized aliquots or 1 mL aliquots when liquid at -80°C until analysis.

Histopathological assessment

Formalin-fixed tissues were embedded overnight and paraffin blocks were prepared. Paraffin sections (1-2 µm) were dewaxed and hydrated in a descending alcohol series. Sections were stained with hematoxylin (Merck) and eosin (Sigma-Aldrich). Sections were dehydrated in an ascending alcohol series with xylene (Carl Roth) as intermediate and coverslipped with corbit balsam (Hecht). Histomorphology of the ileum and colon was evaluated according to modified scores based on Naini and Cortina[1] for CD and Riley[2] for UC. The main modification of both scores include the evaluation of resected tissue with scores for submucosal and transmural inflammation, fissures and increased lymphatic follicles. Minor modifications to the Nini and Cortina scoring system add villous atrophy and fibrosis. Also, for the Riley scoring scheme, the modifications include the scores for resected tissue as well as the scoring for ileal involvement (evaluation of infiltration with acute and chronic inflammatory cells, architectural distortion and epithelial integrity).

Stool score assessment

Stool samples were taken by the patients and shipped in RNAprotect reagent accompanied by a questionnaire. In order to classify various types of feces the Bristol stool chart was used[3].

Whole exome sequencing analysis

Germline mutations were called using a custom-built nextflow pipeline. Briefly: Whole exome sequencing raw reads were cleaned from residual adapter sequences and low-quality sequences using fastp v0.12.4[4]. The reads were then aligned to the reference genome (hg38) using BWA v0.7.17[5]. Duplicate reads were marked with sambamba v0.8.0[6]. Base-call quality score recalibration was performed with GATK4 v4.2.3[7]. Germline variants are called using the haplotypecaller program from GATK4 and Strelka2 v2.9.10[8]. Variants that were called from both algorithms were used as high-confidence variants and annotated using the Ensembl variant effect prediction (VEP v104.3) tool[9].

To investigate *NOD2*, all mutations were filtered to retain only coding variants associated with protein-coding transcripts. Exon regions were extracted from the Gencode v33 primary

assembly annotation GTF file using the R-package GenomicFeatures (v.1.56.0). A transcript database (TxDb) was created with the *makeTxDbFromGFF* function. Transcript names were retrieved using the *transcripts* function and filtered to match *NOD2* transcript IDs present in our dataset. The distribution of *NOD2* mutations was visualized using the trackViewer R-package (v.1.40.0). A lollipop plot was generated, highlighting the most frequent mutations in red.

GenePy score calculation

To estimate the pathogenic burden of genomic variants we adopted the GenePy[10] score. Briefly, the score was calculated from variant effect predictor (VEP) annotations by integrating functional impact, allele frequency, and zygosity for each variant. First, the raw CADD scores[11] (v1.6) for all variants were rescaled between the minimum and maximum values observed across the dataset. To account for population frequency, each variant's gnomAD allele frequency was adjusted to avoid zero or one values, and converted to a probability tuple representing the reference and alternate allele frequencies. The variant-level score was then computed as the scaled CADD score multiplied by the negative log10 of the product of these allele frequencies; for homozygous variants, this value was doubled. Only variants with predicted coding consequences were retained. Finally, variant scores were summed per gene for each individual, producing a gene-level "GenePy" score that reflects the cumulative predicted deleterious burden of coding variants in that gene.

eQTL analysis

To identify cis-expression quantitative trait loci (cis-eQTLs), genotyping was performed using GATK HaplotypeCaller[12] to generate a genomic VCF (GVCF), which was subsequently lifted from GRCh38 to GRCh19 using BCFtools/liftover[13]. Variants were imputed using a local Nextflow pipeline running the Michigan Imputation Server 2[14] in combination with the 1000 Genomes reference panel. Post-imputation quality control was conducted using bcftools[15] (v1.23) and PLINK2[16] (v2.0.0-a.6.9LM); variants were initially filtered for imputation quality ($R^2 > 0.85$) and minor allele frequency (MAF > 0.01). For the eQTL analysis, a stricter subset of variants was generated with thresholds for MAF (> 0.1), genotype missingness (< 0.1), individual missingness (< 0.05), and Hardy-Weinberg equilibrium (hwe $1x10^{-6}$). Gene expression data was processed by filtering low-expression genes (TPM > 0.1 in ≥ 30% of samples) and applying gene-level inverse normal transformation. Hidden confounders were estimated using the PEER[17] (v.2.0) package, with the number of factors (*K*) determined by sample size (*K*=15 for *N*<150, *K*=30 for *N*=150–250, *K*=60 for *N*>250), and combined with optimal genotype principal components (PCs) selected via the elbow method. Cis-eQTL

mapping was performed using MatrixEQTL[18] (v2.3) with a linear model within a 1 Mb window. To identify a single lead SNP per gene, significant associations (FDR < 0.05) were prioritized using a hierarchical tie-breaking strategy: variants were sorted first by lowest FDR, followed by highest empirical imputation quality (ER^2), highest imputation quality (R^2), and finally highest MAF. Fine-mapping was conducted using susieR[19] (v0.14.2) on expression residuals after regressing out covariates to identify 95% credible sets (**Supplementary Table 2**). Significant gene-SNP pairs (FDR < 0.05) were compared with all significant cis-eQTL results from the healthy intestinal datasets of the GTEx[20] project (GTEx eQTL analysis v8: colon transverse, colon sigmoid, small intestine terminal ileum). Colocalization with IBD GWAS summary statistics from the IEU OpenGWAS[21] database (dataset: IBD ebi-a-GCST004131) was assessed using the coloc[22] (v5.2.3) package by requiring a posterior probability (PP4) greater than 0.5 for a model with a single common causal variant and leveraging the effect sizes (β) and their variance (var-β) for shared variants. Within the colocalization model, eQTLs were defined as quantitative traits (type = "quant"), whereas GWAS summary statistics were defined as case-control traits (type = "cc") utilizing the respective sample sizes and case proportions. To assess inflammation-dependent effects, the cohort was stratified into low (≤ 33rd percentile) and high (≥ 66th percentile) inflammation groups based on the normalized histological score. To prevent overcorrection and preserve the stratification signal, this score was excluded from the final covariate matrix during the group-specific eQTL mapping. Finally, context-specific interactions were identified by comparing the resulting group-specific β coefficients using a Z-score test with $p_{interaction} < 0.05$.

Transcriptomics analysis

RNA-sequencing samples from four different batches were processed with the nf-core RNA-seq pipeline version 3.4[23]. In brief, sequencing reads were aligned to the hg38/GRCh38 reference genome with Gencode v33 annotations using STAR v2.7.7a[24]. Read counts and transcripts per million (TPM) were quantified using Salmon[25].

Differential expression analysis was performed in R v.4.4.1 with DESeq2 (v.1.44.0) using raw counts and the covariate formula ~ *group* + *batch* + *sex* + *scaled age*. For comparisons between IBD inflamed and non-IBD samples *tissue_coarse* was added as an additional covariate to account for the different tissues involved. Genes were considered differentially expressed if they met an adjusted p-value threshold of < 0.05 and a |log2FoldChange| threshold of >1. For visualization of the results we used the EnhancedVolcano (v.1.22.0), ggplot2 (v.3.5.1), ComplexHeatmap (v.2.20.0), and ggvenn (v.0.1.10) R-packages.

Cytokine signaling activities for bulk gene expression data were inferred using CytoSig[26] in Python v.3.8.20, leveraging the cytosig.v0.1 implementation available on GitHub (https://github.com/data2intelligence/CytoSig). TPM values were log-transformed as $\log_2(\text{TPM} + 1)$ prior to analysis and used as input. CytoSig calculates the z-score by dividing the regression coefficient by the standard error. The p-values are obtained using a permutation test when the random count is > 0 or using a Student's t-test if the random count is 0.

Biopsy and circulating molecular inflammation signatures were obtained from Argmann et al.[27]. To calculate the biopsy molecular inflammation scores (bMIS) for our samples, we applied gene-set variation analysis (GSVA)[28] using the GSVA R-package (v.1.52.3).

Cell fractions were inferred from TPM-normalized expression values with quanTIseq[29,30] as implemented in the quantiseqr R-package (v.1.12.0). We applied the built-in "TIL10" signature matrix and performed deconvolution with constrained least squares regression. For visualization cell types with no detected fractions (dendritic cells and monocytes) as well as the "Other" category were excluded. Differences between conditions were assessed using the Wilcoxon rank-sum test with significance thresholds set at: $p < 0.05$ (*), $p < 0.01$ (**), $p<0.001$ (***), and $p<0.0001$ (****).

Serum protein analysis

Data tables containing normalized protein expression (NPX) values, Olink Proteomics' arbitrary unit on $\log_2$ scale, were loaded into R v.4.4.1 and further processed with the OlinkAnalyze (v.4.0.1) R-package. Differential protein analysis was conducted using the *olink_ttest* function. Only proteins detected in at least 90% of the measured samples were included in the analysis. Statistical differences were assessed using the Welch two-sample t-test with Benjamini-Hochberg correction applied to adjust for multiple testing. Proteins were considered differentially abundant if they met a FDR threshold of < 0.05. Results were visualized using the EnhancedVolcano (v.1.22.0) R-package. Intersections were retrieved and plotted with the ggVennDiagram (v.1.5.2) or the UpSetR (v.1.4.0) R-package.

We developed an IBD Inflammatory Protein Severity Signature (IBD-IPSS) using a method consistent with the approach outlined by Argmann et al.[27]. In brief, differential protein abundance between inflamed and non-inflamed IBD samples was analyzed using OlinkAnalyze as described above, identifying significantly upregulated proteins for inclusion in the IBD-IPSS. Similarly, entity-specific signatures were generated: the UC-IPSS and CD-IPSS, derived by analyzing protein abundance separately in ulcerative colitis and

Crohn's disease samples. Correlation analysis with the various inflammatory scores available within IBDome including endoscopic scores (SES-CD and UCEIS), clinical scores (HBI and PMS), histopathology scores (modified Riley and modified Naini Cortina score) and the computed bMIS scores (bMIS-CD and bMIS-UC) was conducted using Pearson correlation with pairwise complete observations.

Functional analysis and clustering of the IBD-IPSS proteins was performed using the STRING database[31]. Evidence for protein interactions was considered only from curated databases and experimentally validated interactions. Clustering was performed using MCL (Markov Cluster Algorithm)[32] with an inflation parameter set to 3. Clusters were annotated using the default settings of the STRING database web application. This annotation process prioritized general terms or pathways that summarize multiple specific terms and pathways, derived from various databases integrated within STRING.

Multi-omics factor analysis

For the multi-omics factor analysis, we used MOFA[33] as implemented in the MOFA2 R-package (v.1.14.0). As input for the model, we included four data modalities: genomic, transcriptomic, proteomic, and histological features. The genomic layer consisted of the top 1000 most variable mutations assessed by median absolute deviation of the GenePy-scores. For transcriptomic data, we used gene expression values in log(TPM+1), retaining the top 5000 most variable genes across samples. The proteomic layer included normalized Olink protein expression values (NPX), filtered to retain only proteins detected in at least 10 samples. Finally, the histological features were derived from H&E-stained tissue sections as described in "*Generation of classification-derived feature matrix for MOFA integration*". Samples from the different modalities were merged by subject ID, sampling date, and tissue (total N=1485).

We trained the MOFA2 model using default parameters, except that the *scale_groups* data option was set to *TRUE* and the convergence mode was set to "slow". The model identified 20 latent factors that captured major sources of variation. To visualize the latent structure of the learned factors, we applied Uniform Manifold Approximation and Projection (UMAP) for non-linear dimensionality reduction with the *run_umap* and *plot_dimred* functions as implemented in the MOFA2 R-package.

The IBDome research database

A relational database was designed and implemented in the Python package sqlalchemy using SQLite as database engine. Data integrity is ensured through check constraints and foreign key validation. SQLite was chosen over other database systems, because it makes the database easy to share as a single file, does not require a server, and offers good performance for a use-case without concurrent writes. Inconsistencies in clinical data were resolved manually, and implausible entries were removed. Both clinical and molecular data were processed and imported into the database in a set of Jupyter notebooks and a custom helper library written in Python. All data loading steps are integrated into a Nextflow pipeline[34], which allows rebuilding the database from scratch in a single command.

Web application

The IBDome web application is implemented in R Shiny and directly interacts with the IBDome SQLite database. Dependencies are packaged in a Docker container and a docker-compose file is provided which allows executing the app locally. Plots were generated in R using the ggplot2, ggpubr, plotly, and ggbeeswarm packages. For visualization of gene expression data, transcripts per million (TPM) values were $\log_{10}$(TPM+1) transformed. P-values were computed using a two-tailed Wilcoxon test on the transformed values.

Foundation Model Embedding Extraction

H&E WSI were tessellated into patches with dimensions of 224×224 pixels, representing a 256 µm edge length. To ensure consistent color distribution across cohorts, patches from each cohort underwent color normalization using the Macenko spectral matching technique[35], which maps images to a standardized color space. For performance comparison purposes and to ensure the robustness of our findings, we extracted patch-level feature embeddings using four distinct Foundation models—CHIEF[36], UNI2[37], Virchow2[38,39] and H-optimus-1[40]. These models produced an embedding matrix of dimensions n × 768, n × 1536, n × 2560, and n × 1536, respectively, for each patient's pre-processed patches, where n is the number of (224 ×224 pixels) pre-processed image patches obtained per whole slide image. All preprocessing steps followed the STAMP protocol[41], which harmonizes patch extraction and feature computation across cohorts. The resulting slide-level embedding matrices were subsequently used for both the inflammation score prediction (regression task) and IBD subtype classification (classification task).

Deep Learning Inflammation score prediction

The extracted slide-level embeddings were processed in an attention-based multiple instance learning (attMIL) framework[42] designed for weakly supervised regression. For each foundation model, a separate attMIL model was trained using 5-fold cross-validation on the Berlin cohort to predict the normalized modified Naini Cortina score and the normalized modified Riley score. The cross-validation employed score-based stratification to maintain consistent data distributions across all folds, resulting in five models trained and tested on distinct and balanced splits. To externally validate the model's prognostic performance, all five attMIL models from the cross-validation folds were independently deployed to the Erlangen and Innsbruck cohorts to mitigate fold-specific variability. For the Innsbruck cohort, histologic disease activity scores were assigned independently by averaging the scores of two gastrointestinal histology experts, including one board-certified gastrointestinal pathologist, both blinded to clinical and molecular data, and aggregated by arithmetic averaging to generate a consensus reference score for model evaluation. Slide-level predictions were generated by each model and then aggregated through arithmetic averaging to produce the final prognostic scores. These steps were performed using the open-source Deep Learning pipeline “marugoto”[43], with the default hyperparameters (learning rate = 0.0001, weight decay = 0.01, batch size = 1) using weighted L1 loss.

Explainability of the Deep Learning model

To interpret the decision-making process of the regression models, we leveraged the attention mechanism of the attMIL architecture. High-resolution attention heatmaps were created by loading the attMIL model architectures for regression into a fully convolutional equivalent[44] with their respective weights from the training procedure. By running inference on each patient’s WSI, we extracted the attention layer associated with the score prediction and overlaid it on the WSI, highlighting the regions of focus for the model’s predictions of the scores. For visualization, we selected the Berlin cohort to observe the model performance in predicting disease activity scores. For a more detailed evaluation, we selected the top 10 attention heatmaps for each scoring system based on prediction accuracy. These heatmaps were then reviewed by an expert pathologist, who assessed the highlighted regions for correspondence with areas of known clinical relevance.

IBD subtype classification

In order to evaluate the ability of foundation model embeddings to support subtype discrimination within IBD, we trained a weakly supervised transformer classifier aimed at

distinguishing CD from UC. The classification pipeline was implemented using the STAMP framework[41], utilizing the same slide-level embeddings previously extracted from each foundation model (CHIEF[36], UNI2[37], Virchow2[38,39], and H-optimus-1[40]). For each model, a 5-fold cross-validation was conducted on the Berlin cohort using stratified splitting by disease subtype. The cohort had a near-equal distribution of CD and UC cases, minimizing class imbalance during training. Following cross-validation, a final classifier model was retrained on the entire Berlin cohort and then deployed on four distinct subsets of the Erlangen cohort to assess generalization: all tissue samples (full Erlangen set), all inflamed tissues, colon-only samples, and inflamed colon samples. The classifier was trained using cross-entropy loss, while predictions were obtained by applying a softmax activation to the model outputs, with class assignment based on the highest predicted probability. Model performance metrics were computed for each subset, and 1,000 bootstrap resampling were performed to derive 95% confidence intervals for all performance estimates.

Generation of classification-derived feature matrix for MOFA integration

To derive histopathology features for multi-modal factor analysis, we leveraged the Virchow2 model[39], which showed the strongest classification performance in our benchmarking. A 5-fold cross-validation was performed on the combined Berlin and Erlangen cohorts, and for each image, we used the specific fold model responsible for generating its classification prediction. From the slide-level h5 embedding feature files, we selected the top 10 tiles per slide based on the softmax score for the prediction of the class. A weighted average of the 1280-dimensional embeddings of the top 10 tiles, using the normalized softmax scores as weights, was then computed to obtain a single feature vector per slide. This resulted in a patient-level feature matrix of size 1250 × 1280, which was used as the histology modality input for MOFA[33] integration.

Supplementary Results

Genetic regulation of tissue gene expression

We performed expression quantitative trail locus (eQTL) analysis to examine how genetic variation associates with gene expression in the tissue. cis-eQTL mapping identified 3,600 distinct eGenes. Comparison of significant eSNP-eGene associations with GTEx healthy intestinal eQTL data yielded 3,138 overlapping pairs, including 224 lead cis-eQTLs. Integration of these cis-eQTLs with IBD GWAS summary statistics (EBI study GCST004131) identified 60 shared causal variants with high colocalization probability. Furthermore, our analysis highlighted 396 inflammation-dependent lead eQTLs affecting 388 distinct eGenes (**Supplementary Figure 1B-C, Supplementary Table 1-5)**.

Multi-omic factor analysis reveals tissue- and inflammation-driven variation

To integrate the diverse datasets in our cohort, we performed multi-omic factor analysis using MOFA2[33], incorporating genomics, transcriptomics, serum proteomics, and image-derived features from H&E-stained tissue sections. After merging samples by patient, date, and tissue (**Supplementary Figure 2A**), MOFA identified 20 latent factors explaining variation across modalities (**Supplementary Figure 2B**). The majority of variance was captured by transcriptomics and histology-derived features, followed by proteomics.

The leading factors predominantly captured tissue type and inflammation, consistent with known drivers of molecular heterogeneity in IBD, while disease subtype (CD vs. UC) accounted for only minor differences (**Supplementary Figure 2C**). UMAP embedding showed continuous gradients without clear CD/UC separation (**Supplementary Figure 2D**), highlighting the molecular similarity of inflamed tissues across IBD subtypes and the challenge of distinguishing disease type from bulk tissue data alone.

Supplementary References

1. Naini BV, Cortina G. A histopathologic scoring system as a tool for standardized reporting of chronic (ileo)colitis and independent risk assessment for inflammatory bowel disease. Human Pathology 2012;43:2187–2196.

2. Riley SA, Mani V, Goodman MJ, et al. Microscopic activity in ulcerative colitis: what does it mean? Gut 1991;32:174–178.

3. Lewis SJ, Heaton KW. Stool Form Scale as a Useful Guide to Intestinal Transit Time. Scandinavian Journal of Gastroenterology 1997;32:920–924.

4. Chen S, Zhou Y, Chen Y, et al. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 2018;34:i884–i890.

5. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754–1760.

6. Tarasov A, Vilella AJ, Cuppen E, et al. Sambamba: fast processing of NGS alignment formats. Bioinformatics 2015;31:2032–2034.

7. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics 2013;43:11.10.1-11.10.33.

8. Kim S, Scheffler K, Halpern AL, et al. Strelka2: fast and accurate calling of germline and somatic variants. Nat Methods 2018;15:591–594.

9. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. Genome Biol 2016;17:122.

10. Mossotto E, Ashton JJ, O'Gorman L, et al. GenePy - a score for estimating gene pathogenicity in individuals using next-generation sequencing data. BMC Bioinformatics 2019;20:254.

11. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 2014;46:310–315.

12. Poplin R, Ruano-Rubio V, DePristo MA, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. 2017. Available at: http://biorxiv.org/lookup/doi/10.1101/201178 [Accessed February 18, 2026].

13. Genovese G, Rockweiler NB, Gorman BR, et al. BCFtools/liftover: an accurate and comprehensive tool to convert genetic variants across genome assemblies Kendziorski C, ed. Bioinformatics 2024;40:btae038.

14. **Das S, Forer L, Schönherr S**, et al. Next-generation genotype imputation service and methods. Nat Genet 2016;48:1284–1287.

15. Danecek P, Bonfield JK, Liddle J, et al. Twelve years of SAMtools and BCFtools. Gigascience 2021;10:giab008.

16. Chang CC, Chow CC, Tellier LC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015;4:s13742-015-0047–8.

17. Stegle O, Parts L, Piipari M, et al. Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. Nat Protoc 2012;7:500–507.

18. Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics 2012;28:1353–1358.

19. Cui R, Elzur RA, Kanai M, et al. Improving fine-mapping by modeling infinitesimal effects. Nat Genet 2024;56:162–169.

20. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science 2020;369:1318–1330.

21. Lyon M, Andrews SJ, Elsworth B, et al. The variant call format provides efficient and robust storage of GWAS summary statistics. 2020. Available at: http://biorxiv.org/lookup/doi/10.1101/2020.05.29.115824 [Accessed February 23, 2026].

22. Wallace C. A more accurate method for colocalisation analysis allowing for multiple causal variants Cordell HJ, ed. PLoS Genet 2021;17:e1009440.

23. **Ewels PA, Peltzer A**, Fillinger S, et al. The nf-core framework for community-curated bioinformatics pipelines. Nat Biotechnol 2020;38:276–278.

24. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15–21.

25. Patro R, Duggal G, Love MI, et al. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 2017;14:417–419.

26. **Jiang P, Zhang Y, Ru B**, et al. Systematic investigation of cytokine signaling activity at the tissue and single-cell levels. Nat Methods 2021;18:1181–1191.

27. Argmann C, Hou R, Ungaro RC, et al. Biopsy and blood-based molecular biomarker of inflammation in IBD. Gut 2023;72:1271–1287.

28. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 2013;14:7.

29. Finotello F, Mayer C, Plattner C, et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. Genome Med 2019;11:34.

30. Plattner C, Finotello F, Rieder D. Chapter Ten - Deconvoluting tumor-infiltrating immune cells from RNA-seq data using quanTIseq. In: Galluzzi L, Rudqvist N-P, eds. *Methods in Enzymology*.Vol 636. Tumor Immunology and Immunotherapy – Integrated Methods Part B. Academic Press; 2020:261–285. Available at: https://www.sciencedirect.com/science/article/pii/S007668791930237X [Accessed August 11, 2025].

31. Szklarczyk D, Kirsch R, Koutrouli M, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res 2023;51:D638–D646.

32. Van Dongen S. Graph Clustering Via a Discrete Uncoupling Process. SIAM J Matrix Anal Appl 2008;30:121–141.

33. **Argelaguet R, Velten B**, Arnol D, et al. Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. Molecular Systems Biology 2018;14:e8124.

34. Di Tommaso P, Chatzou M, Floden EW, et al. Nextflow enables reproducible computational workflows. Nat Biotechnol 2017;35:316–319.

35. Macenko M, Niethammer M, Marron JS, et al. A method for normalizing histology slides for quantitative analysis. In: *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*.; 2009:1107–1110. Available at: https://ieeexplore.ieee.org/document/5193250 [Accessed October 30, 2024].

36. **Wang X, Zhao J**, Marostica E, et al. A pathology foundation model for cancer diagnosis and prognosis prediction. Nature 2024;634:970–978.

37. **Chen RJ, Ding T, Lu MY,** et al. Towards a general-purpose foundation model for computational pathology. Nat Med 2024;30:850–862.

38. **Vorontsov E, Bozkurt A, Casson A**, et al. A foundation model for clinical-grade computational pathology and rare cancers detection. Nat Med 2024;30:2924–2935.

39. Zimmermann E, Vorontsov E, Viret J, et al. Virchow2: Scaling Self-Supervised Mixed Magnification Models in Pathology. arXiv e-prints 2024:arXiv:2408.00738.

40. Saillard C, Jenatton R, Llinares-Lopez F, et al. H-optimus-1. 2024. Available at: https://github.com/bioptimus/releases/tree/main/models/h-optimus/v0.

41. El Nahhas OSM, Treeck M van, Wölflein G, et al. From whole-slide image to biomarker prediction: end-to-end weakly supervised deep learning in computational pathology. Nat Protoc 2024:1–24.

42. Leiby JS, Hao J, Kang GH, et al. Attention-based multiple instance learning with self-supervision to predict microsatellite instability in colorectal cancer from histology whole-slide images. In: *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*.; 2022:3068–3071. Available at: https://ieeexplore.ieee.org/document/9871553 [Accessed October 30, 2024].

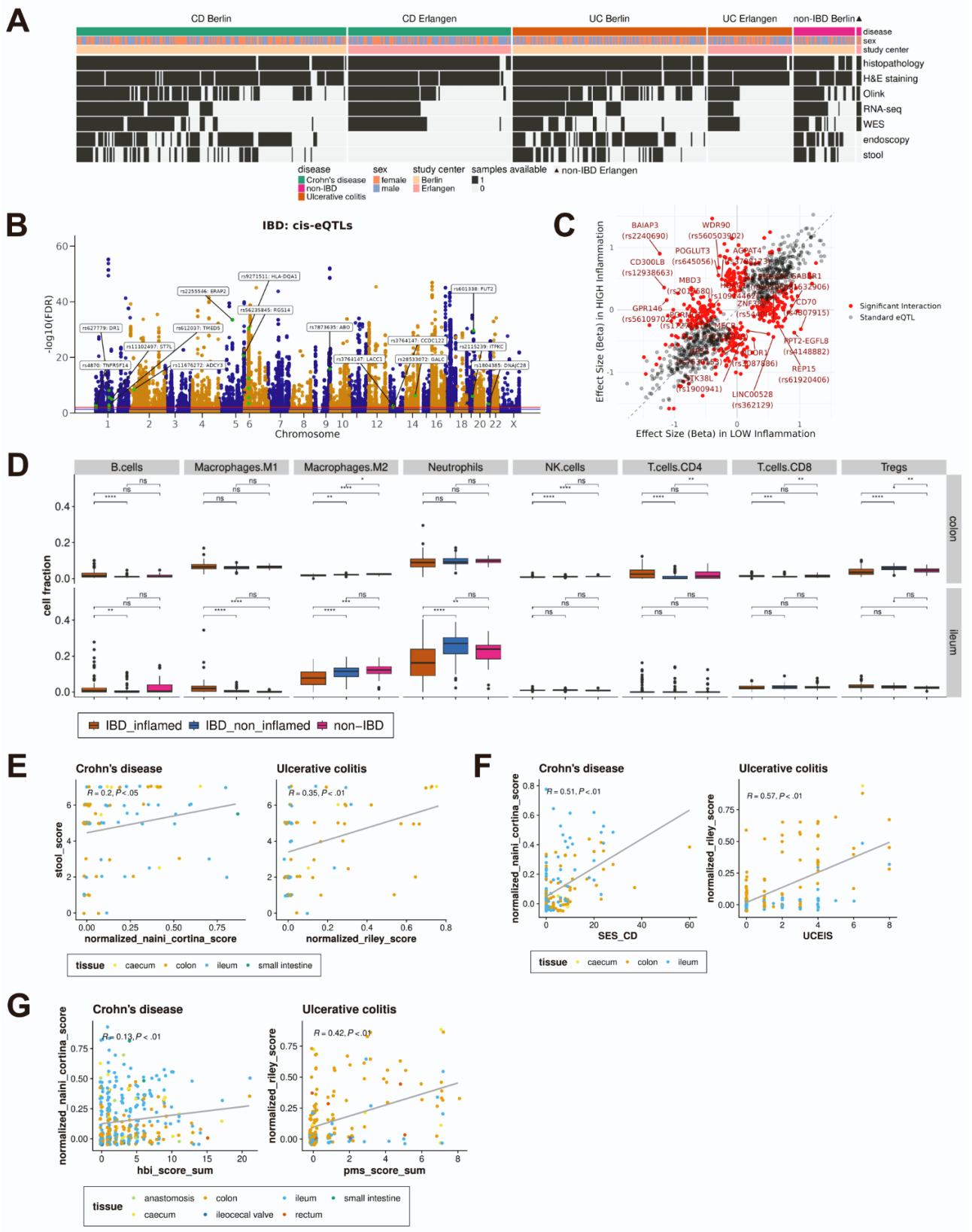
43. Ilse M, Tomczak J, Welling M. Attention-based Deep Multiple Instance Learning. In: *Proceedings of the 35th International Conference on Machine Learning*. PMLR; 2018:2127–2136. Available at: https://proceedings.mlr.press/v80/ilse18a.html [Accessed October 30, 2024].

44. Pathak D, Shelhamer E, Long J, et al. Fully Convolutional Multi-Class Multiple Instance Learning. 2015. Available at: http://arxiv.org/abs/1412.7144 [Accessed October 30, 2024].

Author names in bold designate shared co-first authorship.

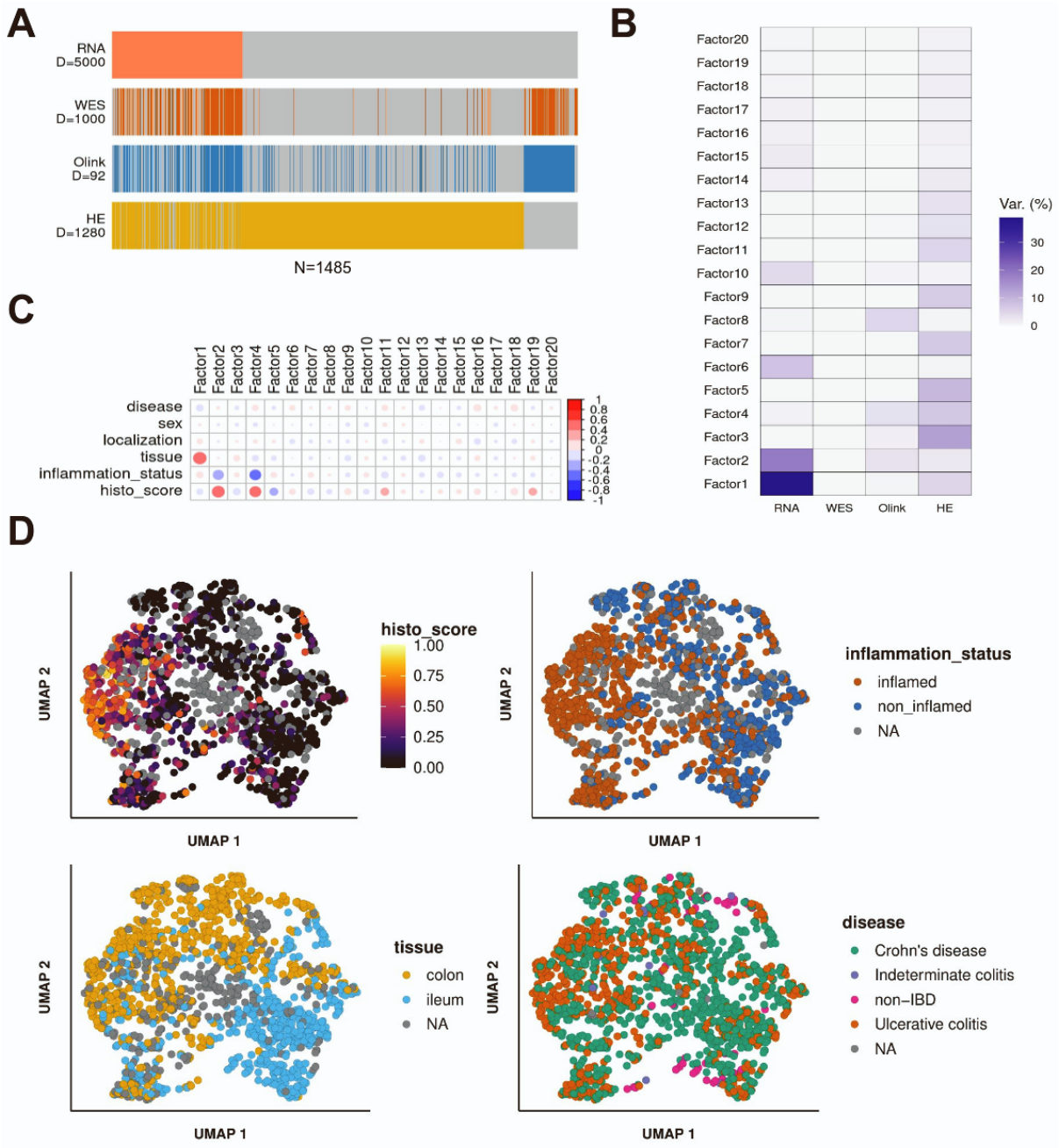
Supplementary Figures

Supplementary Figure 1



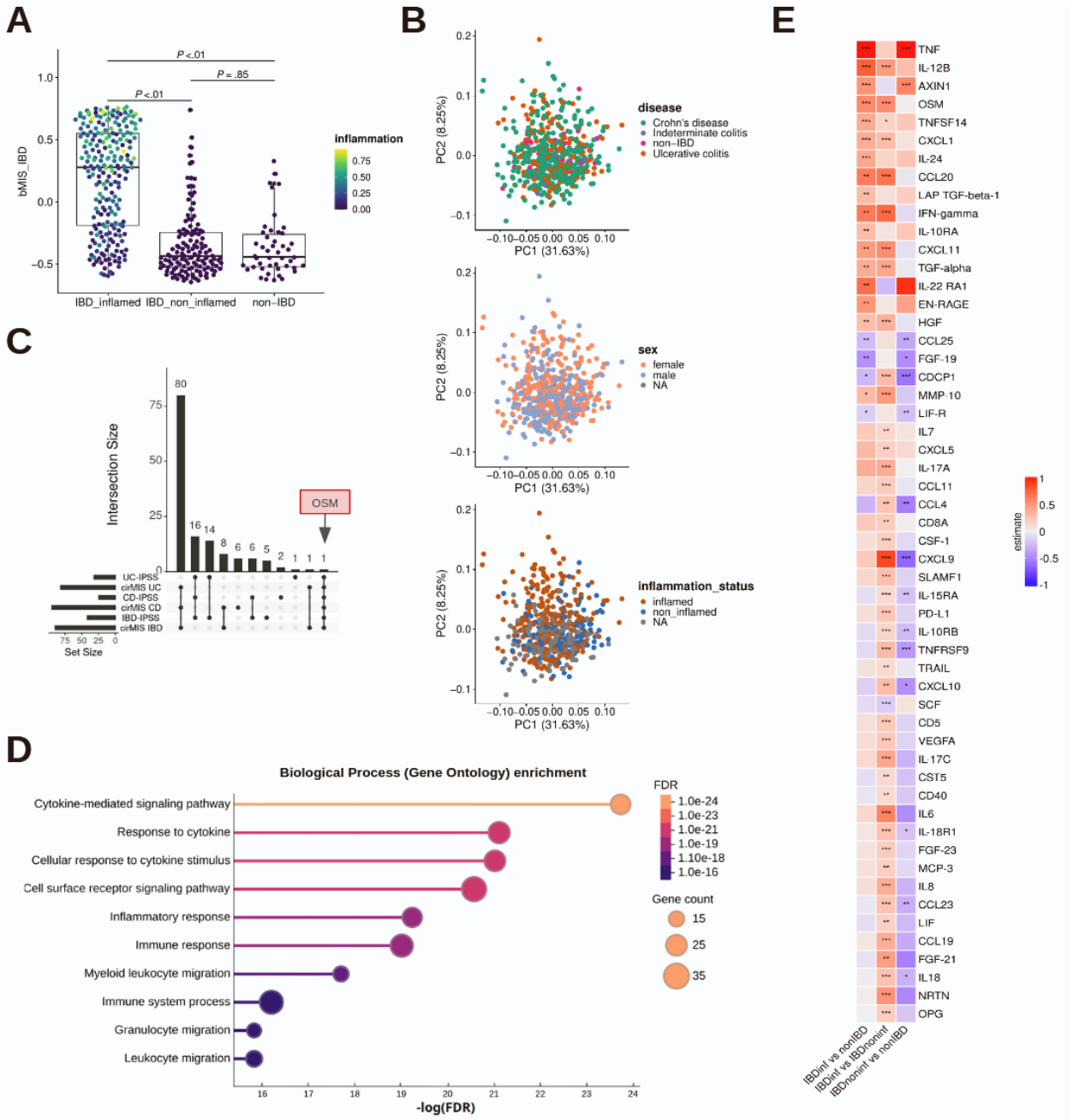
Supplementary Figure 1. IBDome overview. (A) Cohort distribution illustrated in a heatmap showing the available samples per patient. The heatmap is split by disease and study center and ordered according to the amount of different sample types available per patient. **(B)** Manhattan plot of cis-eQTL associations in the IBDome cohort. The x-axis displays the genomic position of tested variants across chromosomes, and the y-axis represents the significance of the eQTL associations expressed as -log10(FDR). The solid horizontal lines denote False Discovery Rate (FDR) significance thresholds at 0.05 (blue line) and 0.01 (red line). Highlighted data points (green) indicate the cis-eQTLs identified as shared causal variants with IBD susceptibility loci, determined through colocalization analysis with IBD GWAS data (ebi-a-GCST004131). The labels specify the top 15 lead eSNP identifiers and their corresponding eGene. **(C)** Comparison of genetic effect sizes β for inflammation-dependent cis-eQTLs. Low inflammation group (x-axis) versus the high inflammation group (y-axis). Red data points indicate cis-eQTLs with a significant inflammation-dependent interaction ($p_{interaction} < 0.05$). Grey points represent standard eQTLs that do not demonstrate a significant context-dependent interaction. **(D)** Immune cell fractions estimated with quanTIseq with p-values assessed by Wilcoxon test with * $P < .05$, ** $P < .01$, *** $P < .001$, **** $P < .0001$. **(E)** Correlation between histopathology (normalized Riley score and normalized Naini-Cortina score) and the Bristol stool score. **(F)** Correlation between endoscopic (UCEIS = Ulcerative Colitis Endoscopic Index of Severity, SES-CD = Simple Endoscopic Score for Crohn's Disease) and histopathology scores for UC and CD, respectively. **(G)** Correlation between histopathology (normalized Riley score and normalized Naini-Cortina score) and clinical activity scores (PMS= Partial Mayo Score, HBI=Harvey-Bradshaw Index) for UC and CD, respectively.

Supplementary Figure 2



Supplementary Figure 2. Multi-omics factor analysis. (A) Overview of the cohort merged by subject ID, date, and tissue across the four input modalities: RNA-seq (log(TPM+1), top 5000 most variable genes), whole-exome sequencing (WES, GenePy-score, top 1000 most variable mutations), serum proteomics (Olink), and image-derived features from H&E stained tissue sections (HE). **(B)** Variance explained (%) by each MOFA factor per modality. **(C)** Pearson correlation of MOFA factors with clinical covariates. **(D)** UMAP embedding of MOFA factors colored by the normalized histopathology score, inflammation status, tissue type, and disease subtype.

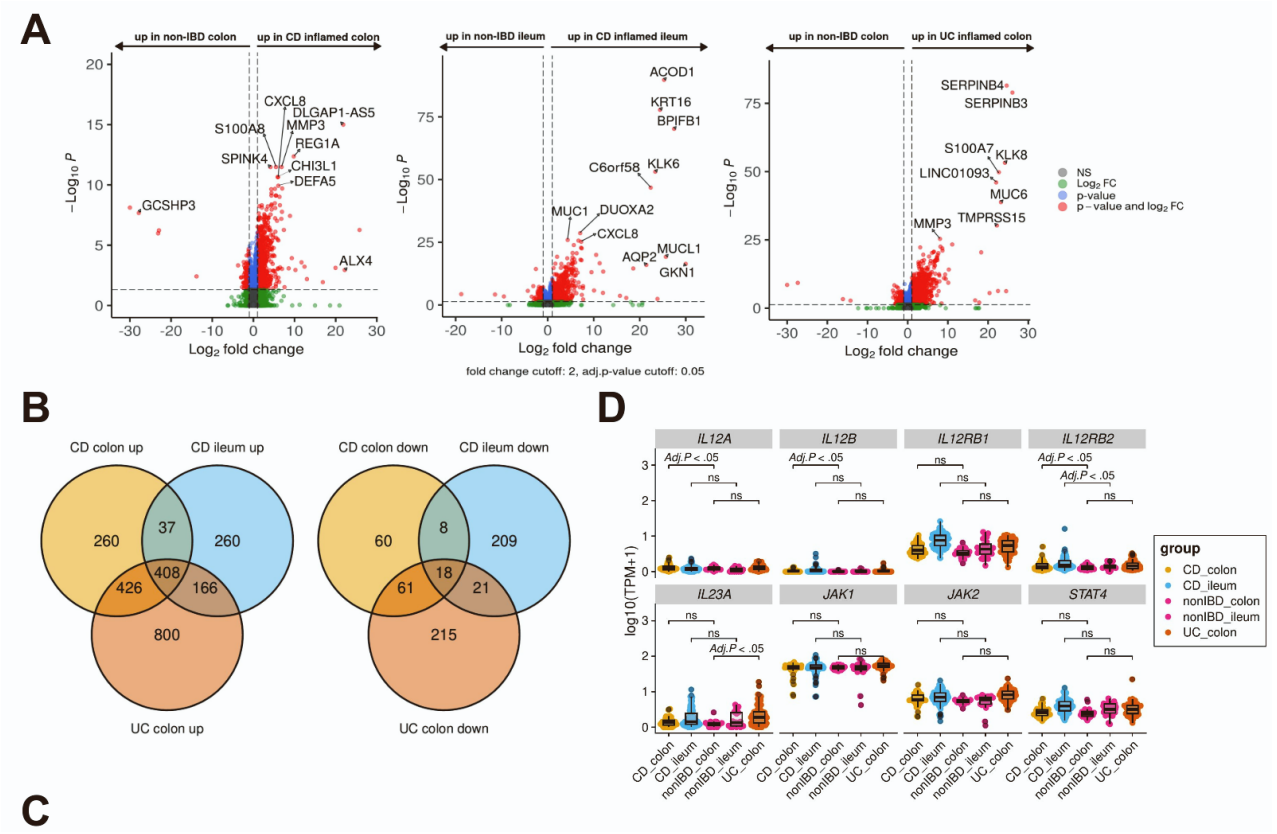
Supplementary Figure 3



Supplementary Figure 3. Protein abundance in the serum. (A) Biopsy-based molecular inflammation scores (bMIS-IBD) across different groups, calculated using Gene Set Variation Analysis (GSVA); p-values were assessed using the Wilcoxon test. The color depicts the normalized inflammation score assessed by histopathology (modified Riley score for UC and modified Naini Cortina score for CD). **(B)** Principal Component Analysis (PCA) of serum protein abundances, colored by disease type, sex, and inflammation status **(C)** UpSet plot showing the intersection of different blood-based scores (cirMIS = circulating molecular inflammation score from RNA-seq from the blood, IPSS = inflammatory protein severity

signature from serum proteins) highlighting OSM as the only protein shared across all signatures. **(D)** Significantly enriched Gene Ontology - Biological Processes (GO-BP) terms (FDR < 0.05) from over representation analysis (ORA) in the IBD-IPSS. **(E)** Heatmap of the differentially abundant proteins with significance levels determined by t-test: *** = adj.p<0.01, ** = adj.p < 0.05, * = adj.p < 0.1; “estimate” represents the numeric difference in mean NPX (normalized protein expression; Olink’s arbitrary unit) between groups.

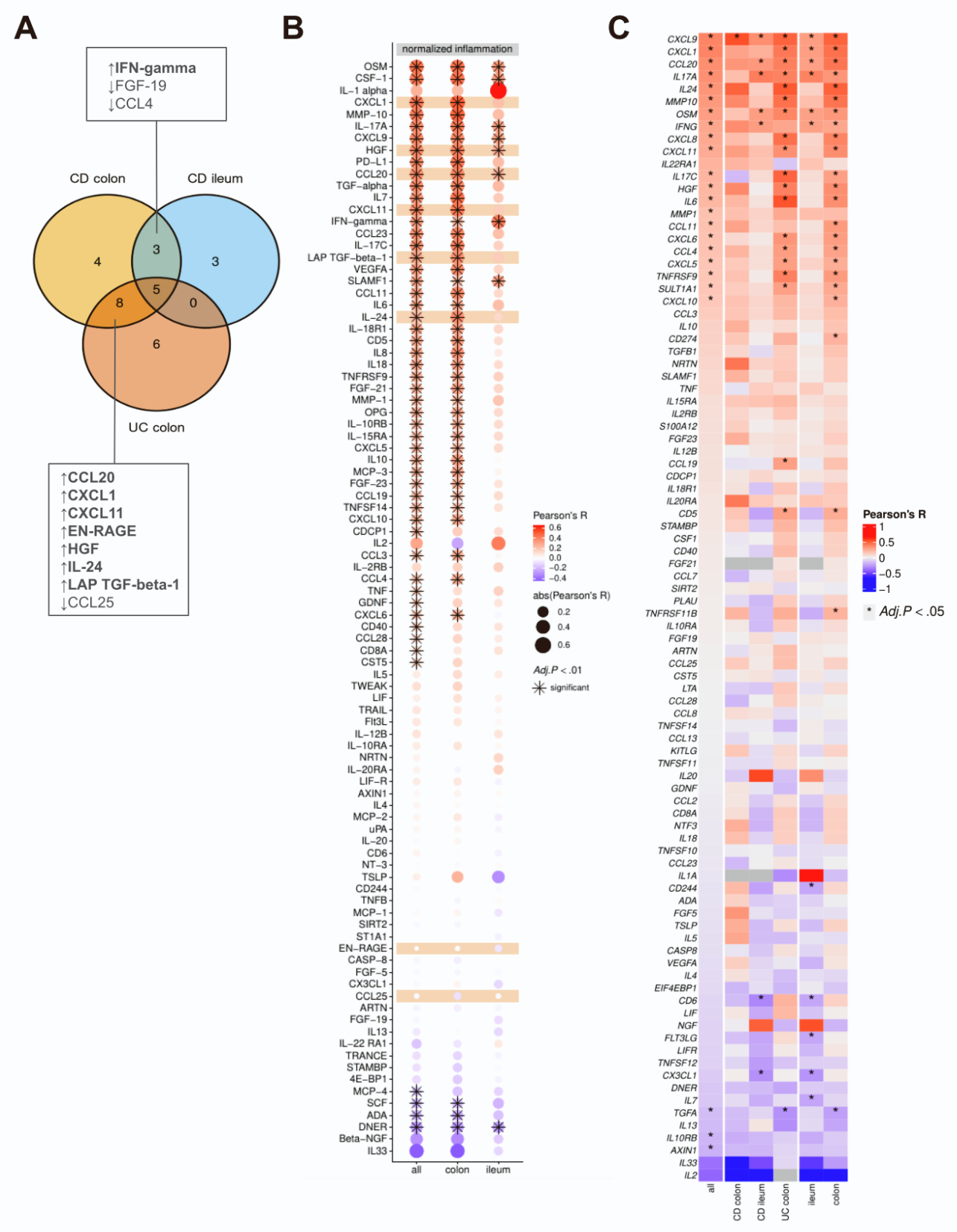
Supplementary Figure 4



Supplementary Figure 4. Transcriptional characterization. (A) Volcano plots displaying the DE analysis results of inflamed CD vs. non-IBD colon samples, inflamed CD vs. non-IBD ileum samples and inflamed UC colon vs. non-IBD colon samples with thresholds: |log2FC|>1 and adjusted *P*<.05. **(B)** Overlap of significantly upregulated and downregulated

genes in the different comparisons. **(C)** Heatmap displaying the gene expression of the commonly downregulated genes (n=18) across the different groups (inflamed IBD, non-inflamed IBD and non-IBD) clustered by euclidean distance and complete linkage. **(D)** Gene expression of genes involved in the IL12 signaling pathway with adjusted p-values retrieved from the differential expression analysis with DESeq2.

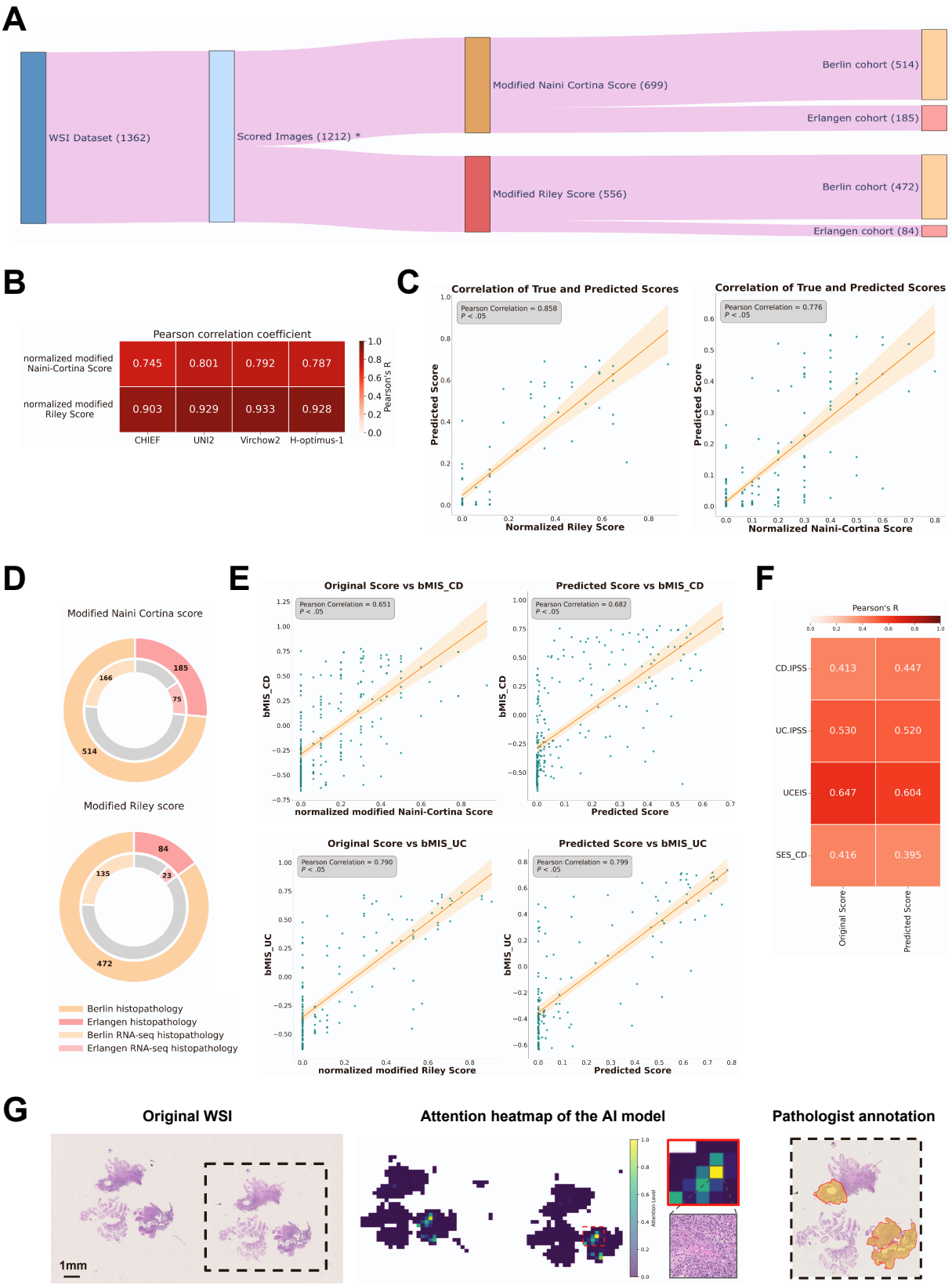
Supplementary Figure 5



Supplementary Figure 5. Correlation of serum proteins. (A) Venn diagram depicting the intersections of differential abundant proteins in the different comparisons. **(B)** Dot plot of

Pearson correlation coefficients (R) between serum protein abundance and histopathology scores (normalized inflammation = normalized modified Riley or normalized modified Naini Cortina score) across all samples and the tissue groups; adjusted p-value cutoff: 0.01; **(C)** Heatmap of Pearson correlation coefficients between serum protein abundance and tissue gene expression across all samples (n=335) and splitted by group and tissue; * adjusted p-value <0.05.

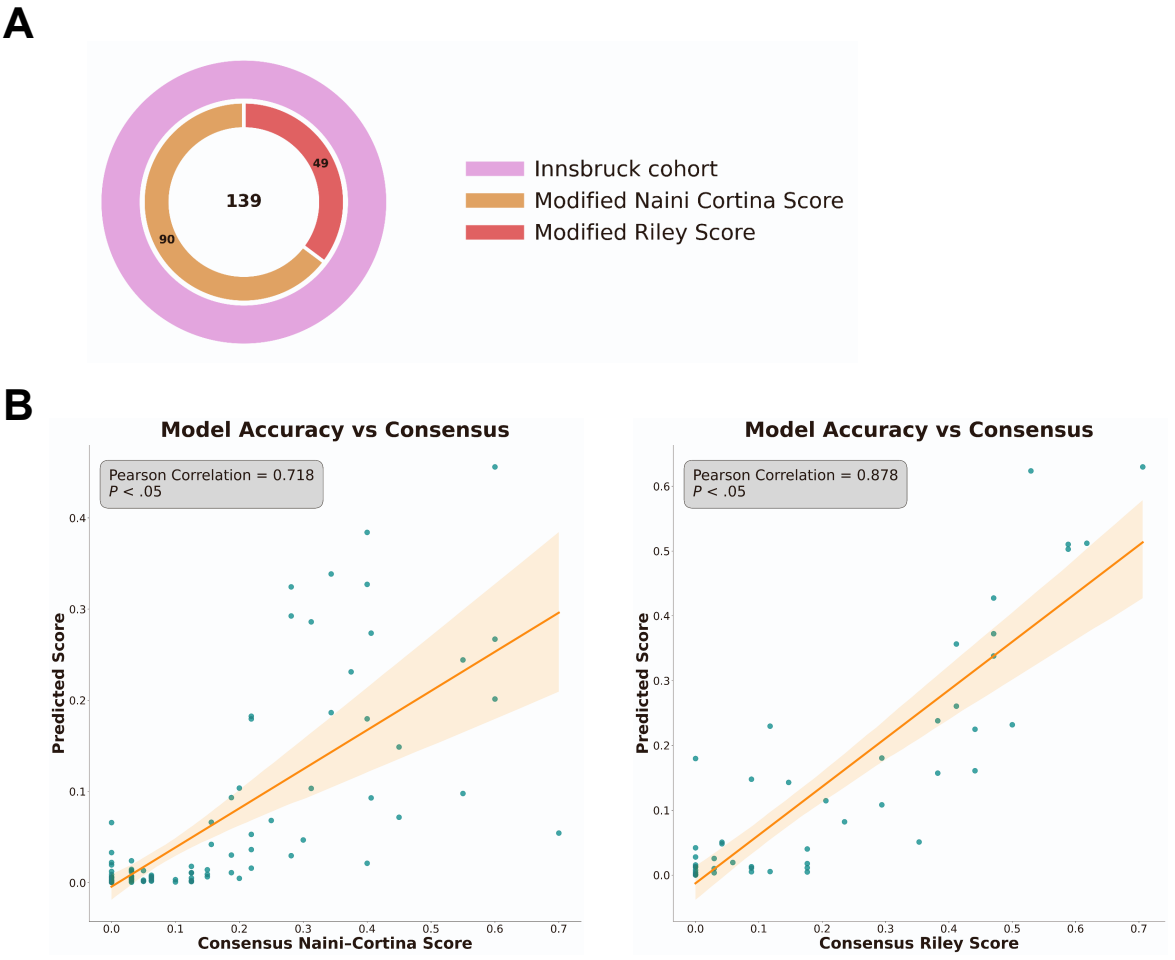
Supplementary Figure 6



Supplementary Figure 6. Prediction of histopathologic scores from H&E images. (A)

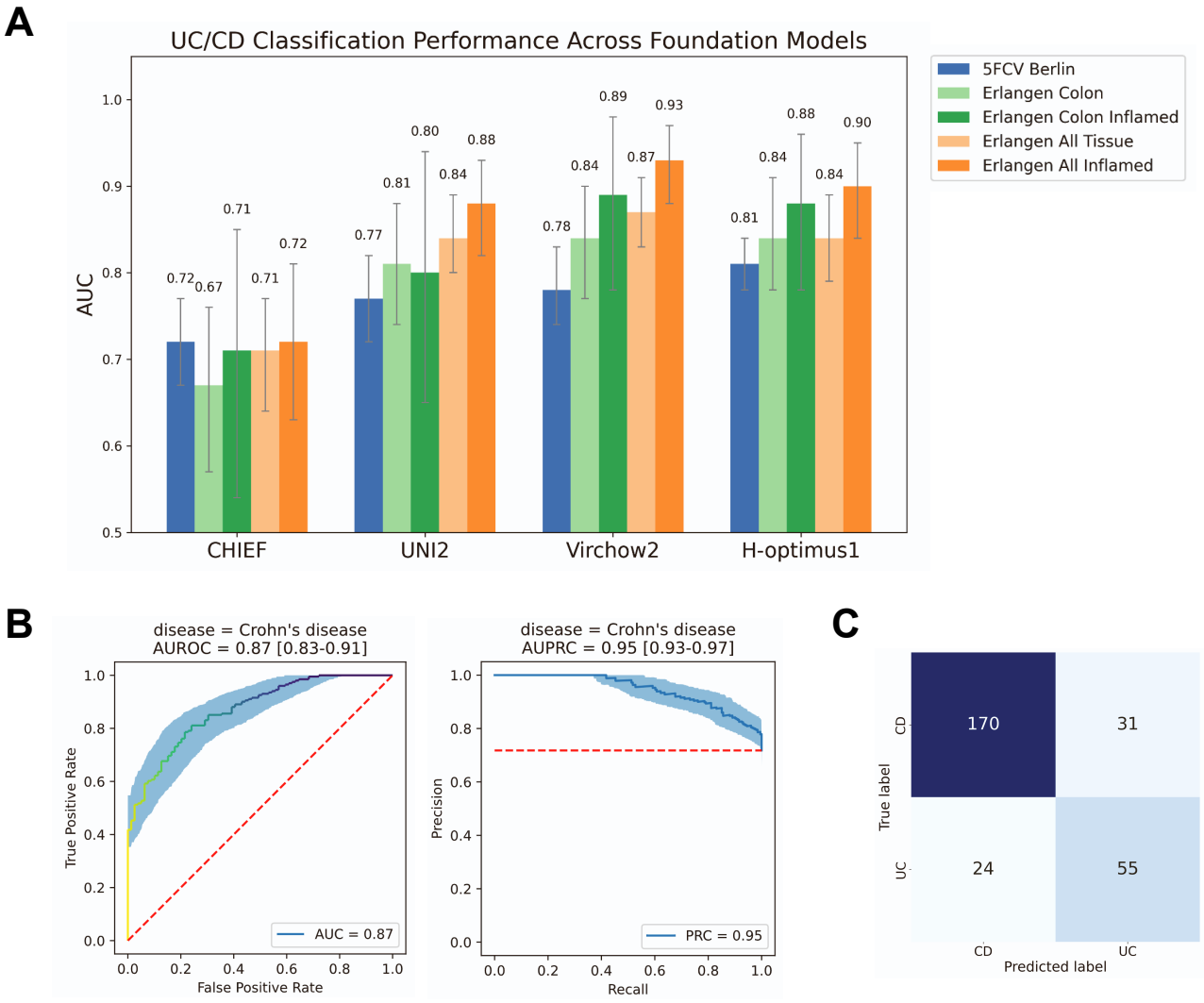
Distribution of all H&E images (n = 1362). After preprocessing, the number of images is reduced to 1212*, which are then categorized with two scoring systems: the modified Naini Cortina score (n = 699) and the modified Riley score (n = 556). Each scoring category is further divided in the two center subsets: Berlin (n = 514 for modified Naini Cortina, n = 472 for modified Riley) and Erlangen (n = 185 for modified Naini Cortina, n = 84 for modified Riley). *Some indeterminate colitis and non-IBD samples were scored with both scoring systems. **(B)** Performance of different Foundation Models (CHIEF, UNI2, Virchow2 and H-optimus-1) on the 5 folds cross-validation regression task on the Berlin cohort, shown by Pearson correlation (R) coefficients. **(C)** Correlation plots between the original histologic disease activity scores (x-axis) and AI-predicted scores (y-axis) for both modified Naini Cortina and modified Riley scoring systems. Predictions are based on the ensemble of 5 cross-validation models trained on the Berlin cohort and evaluated on the Erlangen cohort. **(D)** Cohort distribution: the outer circles represent the Berlin and Erlangen histopathology cohorts, and the inner circles indicate the proportion of available RNA-seq data within each cohort. **(E)** Correlation plots between the bMIS score for UC and CD with the original score (normalized Naini Cortina and Riley score) on the left and the model's predicted scores on the right. **(F)** Comparison of Pearson correlation coefficients (R) between original and predicted histologic disease activity scores against CD-IPSS, UC-IPSS, UCEIS, and SES-CD. **(G)** On the left side, a representative attention heatmap of a biopsy slide image from a CD patient with high histologic disease activity. The heatmap shows the model's attention levels. Higher scores (yellow) mark regions that strongly influence the model's prediction, while lower scores (green) indicate less critical regions. In the middle, a zoomed-in view of the highest-attention region showing the top attention tile and on the right side the pathologist annotation highlighting the important areas to consider for disease activity scoring.

Supplementary Figure 7



Supplementary Figure 7. Histopathologic score prediction from H&E images in the validation cohort. (A) Cohort distribution: The outer circle shows the Innsbruck cohort (n=139), and the inner circle indicates the proportion of modified Naini-Cortina score (n = 90) and modified Riley score (n=49) samples. **(B)** Correlation plots between the consensus histologic disease activity scores (x-axis) and AI-predicted scores (y-axis) for both modified Naini Cortina and modified Riley scoring systems. Predictions are based on the ensemble of 5 cross-validation models trained on the Berlin cohort and evaluated on the Innsbruck cohort.

Supplementary Figure 8



Supplementary Figure 8. Prediction of disease type from pathology images. (A) Classification performance of four Foundation Models (CHIEF, UNI2, Virchow2 and H-optimus-1) across five evaluation settings: 5-fold cross-validation on the Berlin cohort, and training on the same cohort with deployment on the Erlangen cohort using either colon samples, colon-inflamed samples, all tissue types, or all inflamed tissues. Bars represent AUROC values, with error bars denoting 95% confidence intervals. **(B)** AUROC and AUPRC curves for the best-performing model (Virchow2) trained on the Berlin cohort and evaluated on the Erlangen all-tissue subset. **(C)** Confusion matrix summarizing the UC/CD classification results from the model deployment in (B).