




OPEN ACCESS

TRANSLATIONAL SCIENCE

Lupus and inflammatory bowel disease share a common set of microbiome features distinct from other autoimmune disorders

Hao Zhou,¹ Diana Balint,¹ Qiaojuan Shi,¹ Tim Vartanian,² Martin A Kriegel,^{3,4,5,6} Ilana Brito ¹

Handling editor Josef S Smolen

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/ard-2024-225829>).

¹Meinig School of Biomedical Engineering, Cornell University, Ithaca, New York, USA

²Weill Cornell Medicine, New York, New York, USA

³Department of Translational Rheumatology and Immunology, Institute of Musculoskeletal Medicine, Münster, Germany

⁴Section of Rheumatology and Clinical Immunology, University Hospital Münster, Münster, Germany

⁵Cells in Motion Interfaculty Centre, University of Münster, Münster, Germany

⁶Department of Immunobiology, Yale School of Medicine, New Haven, Connecticut, USA

Correspondence to

Dr Ilana Brito;
ibrito@cornell.edu

HZ and DB contributed equally.

Received 20 March 2024
Accepted 3 September 2024



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ on behalf of EULAR.

To cite: Zhou H, Balint D, Shi Q, et al. *Ann Rheum Dis* Epub ahead of print: [please include Day Month Year]. doi:10.1136/ard-2024-225829

ABSTRACT

Objectives This study aims to elucidate the microbial signatures associated with autoimmune diseases, particularly systemic lupus erythematosus (SLE) and inflammatory bowel disease (IBD), compared with colorectal cancer (CRC), to identify unique biomarkers and shared microbial mechanisms that could inform specific treatment protocols.

Methods We analysed metagenomic datasets from patient cohorts with six autoimmune conditions—SLE, IBD, multiple sclerosis, myasthenia gravis, Graves' disease and ankylosing spondylitis—contrasting these with CRC metagenomes to delineate disease-specific microbial profiles. The study focused on identifying predictive biomarkers from species profiles and functional genes, integrating protein-protein interaction analyses to explore effector-like proteins and their targets in key signalling pathways.

Results Distinct microbial signatures were identified across autoimmune disorders, with notable overlaps between SLE and IBD, suggesting shared microbial underpinnings. Significant predictive biomarkers highlighted the diverse microbial influences across these conditions. Protein-protein interaction analyses revealed interactions targeting glucocorticoid signalling, antigen presentation and interleukin-12 signalling pathways, offering insights into possible common disease mechanisms. Experimental validation confirmed interactions between the host protein glucocorticoid receptor (NR3C1) and specific gut bacteria-derived proteins, which may have therapeutic implications for inflammatory disorders like SLE and IBD.

Conclusions Our findings underscore the gut microbiome's critical role in autoimmune diseases, offering insights into shared and distinct microbial signatures. The study highlights the potential importance of microbial biomarkers in understanding disease mechanisms and guiding treatment strategies, paving the way for novel therapeutic approaches based on microbial profiles.

Trial registration number NCT02394964.

INTRODUCTION

Mounting evidence links the gut microbiome with a wide range of autoimmune conditions, including Crohn's disease (CD), ulcerative colitis (UC), rheumatoid arthritis (RA), type 1 diabetes (T1D), multiple sclerosis (MS), systemic lupus

WHAT IS ALREADY KNOWN ON THIS TOPIC

- ⇒ The gut microbiota acts causally in autoimmune diseases, with certain microbial species linked to specific conditions.
- ⇒ There is a significant need for comprehensive analyses to identify biomarkers and understand the mechanisms through which the microbiome influences autoimmune disorders.

WHAT THIS STUDY ADDS

- ⇒ Our research uncovered unique microbial signatures and functional gene profiles across various autoimmune diseases, indicating distinct microbial influences and shared mechanisms, notably between systemic lupus erythematosus and inflammatory bowel disease.
- ⇒ We introduced novel insights through protein-protein interaction analyses, uncovering microbiome-driven pathways that may be crucial for understanding and managing autoimmune conditions.
- ⇒ Additionally, we experimentally validated several key protein-protein interactions, particularly those involving the glucocorticoid receptor (NR3C1), reinforcing their potential as therapeutic targets.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- ⇒ The findings set the stage for future investigations into the interactions between the gut microbiome and the immune system, pointing to potential therapeutic targets.
- ⇒ In clinical settings, the identified microbial biomarkers might facilitate improved diagnosis and personalised treatment strategies for autoimmune diseases, impacting healthcare policies and guiding research towards incorporating metagenomic data for enhanced patient care.

erythematosus (SLE) and others.^{1–9} Metagenomic data from human gut microbiome samples has been used to identify microbial signatures that associate with a particular condition or disease with the goal of identifying potential causative agents and therapeutic targets.^{10–13} Given the overwhelming amount of sequencing data generated, large-scale

comparative analyses can now be used to reliably identify associations, either at the level of species or functions, between gut microbiota and autoimmune diseases and unveil novel biological mechanisms underlying these diseases.

The gut microbiome in autoimmune diseases exhibits significant instability, known as dysbiosis, which manifests distinctly across individuals, being notably pronounced during active phases of the disease and often mild or absent during periods of remission. Such variability suggests that the microbiome's composition is closely linked to the clinical manifestations and the waxing and waning course of autoimmune conditions.^{14 15} This insight has important implications for the use of microbiome profiling as a potential biomarker for disease status and activity such as microbial reactivity,⁹ as well as disease management decisions. Furthermore, the dynamic nature of the microbiome in response to disease activity highlights the necessity for longitudinal studies to capture these fluctuations, thereby enriching our understanding of the microbiome's role in health and disease.^{8 16 17}

Certain microbial species have long been associated with specific conditions or diseases, including in patients with autoimmune diseases, such as SLE and RA.^{6 8 18–21} Mechanistic preclinical studies together with associations in human tissues have indicated that *Enterococcus gallinarum* may be involved in the pathogenesis of SLE and autoimmune liver diseases by translocating into internal organs and promoting self-directed responses.^{6 9 22–25} Similarly, lactobacilli, increased in a subset of patients with SLE, were shown in murine studies to translocate into extraintestinal organs where they promote type I interferon signatures.²¹ Meanwhile, *Ruminococcus gnavus* expansion has been observed in patients with lupus nephritis. Its significance lies in its unique immunoreactive lipoglycan,²⁶ which has immunological implications in lupus.¹⁷ By contrast, higher abundances of *Clostridium leptum*, *Lactobacillus gasseri* and *Bifidobacterium bifidum* are detected in the gut microbiomes of healthy patients, suggesting that these related bacteria may protect against autoimmune disease.^{6 17 27 28} Nevertheless, identifying bacterial species signatures that are specific to a particular autoimmune disease can be challenging due to significant overlap between diseases, especially in the early stages of a disease when the microbiome is in a state of transition.^{29 30} Furthermore, species-level associations provide little mechanistic insight and do not account for the differences in genomic content and functional profiles of individual strains.

Recent metagenomic studies have further advanced our understanding of the microbiome's role in autoimmune diseases, by detailing species and functional capacities that differ in disease states, including involvement in metabolic pathways and drug resistance.³¹ *crp* (K10914, CRP/FNR family transcriptional regulator) has been identified as an important feature of microbiomes in patients with inflammatory bowel disease (IBD) and is associated with higher calprotectin, a measure of gut inflammation.³² Other IBD studies point to enzymes involved in oxidative stress, lipid metabolism and protein degradation.^{13 33} Fewer metagenomic studies have been performed on other autoimmune diseases compared with IBD. However, studies of the gut metagenomes in patients with SLE have found enrichment of pathways related to flagellar assembly, sulfur metabolism and lipopolysaccharide biosynthesis.^{6 34} Metagenomics has the potential for elucidating the microbiome's multifaceted impact on autoimmune pathogenesis,^{6 35} yet more work is needed to establish causality, as well as to establish specificity to each disease.

Host-microbiome protein-protein interactions (PPIs) potentially play critical roles in generating disease-specific signatures

and shaping the unique microbiome profiles observed across different patients in a given disease group.^{3 8} Physical host-microbiota interaction are facilitated by the direct localisation of bacteria in mesenteric lymph nodes, Peyer's patches and epithelial barriers that promote mutualism through both innate and adaptive immune responses.^{36 37} This physical proximity facilitates both mutualistic relationships, and intricate molecular interactions that can impact disease progression. Commensal proteins that include mucin degradation enzymes and protease inhibitors produced by microbiota may have immunomodulatory effects in the intestine that influence immune activation or suppress expression of inflammatory cytokines.^{38 39} To predict microbial functions associated with the development of each autoimmune disease, we used previously constructed host-microbiome PPI network models.⁴⁰ Beyond the observation of a general dysbiosis, we aim to obtain functional insights into species-level differences, involving analyses of differentially abundant PPIs relevant to IBD and/or SLE.

METHODS

Sample collection and processing

78 faecal samples were obtained, with 32 samples from 14 SLE pretreatment outpatients and 46 samples from 22 healthy controls collected at Yale University Medical. The study encompassed patients with lupus and age-matched (± 5 years), as well as sex-matched healthy controls, all of whom were enrolled over a span of 2 years.⁸ The SLE Disease Activity Index (SLEDAI) score was determined using the SLEDAI-2K calculator.⁴¹ Each participant attended up to three study visits, during which comprehensive health and diet histories, whole blood and oral, skin and faecal microbiota samples were collected, in line with a published microbiome study protocol at Yale (ClinicalTrials.gov ID NCT02394964). Faecal samples were then preserved at -80°C pending DNA extraction. As previously mentioned in an earlier study, a dataset including samples from individuals with MS (25 cases and 26 controls, one sample per individual) was collected and incorporated into the analysis.⁴² In the study by Chen *et al* (SLE2), samples from treated patients were excluded from the analysis.¹⁸

Metagenomic sequencing

Genomic DNA was extracted from thawed stool samples using the QIAGEN DNeasy 96 PowerSoil Pro Kit. The DNA samples were then diluted to 0.2 ng/ μL in nuclease-free water and the libraries were prepared using Nextera XT DNA Library Prep Kit (Illumina). Libraries were purified with Ampure XP beads (Beckman Coulter). High-throughput sequencing was carried out by The Yale Center for Genome Analysis using an Illumina NovaSeq6000 system.

Processing of metagenomic data

Sequencing reads were dereplicated using the prinseq-lite.pl V.0.20.26, with the following settings: -derep 12345 -no_qual_header.⁴³ Dereplicated reads were then passed through the KneadData V.0.3 quality control pipeline (<http://huttenhower.sph.harvard.edu/kneaddata>), which incorporates the Trimmomatic and BMTagger⁴⁴ and decontamination algorithms to remove low-quality reads (thresholding Phred quality score at <20 ; minimum length <150) and reads of human genome origin.⁴⁵ Taxonomic profiling was performed using MetaPhlAn3. Functional profiling was performed using HUMANN3.⁴⁶ Samples with fewer than 10^7 reads were removed from analyses. Short-chain fatty acids (SCFAs) were identified and concatenated

according to Kyoto Encyclopedia of Genes and Genomes (KEGG) orthologs, as catalogued in a prior study.⁴⁷

CRC and IBD datasets used in this study were curated as part of ExperimentHub.⁴⁸ We downloaded all protein abundance matrices, annotated at the level of UniRef90 clusters via HUMAnN3, and associated metadata. For each study, we mapped UniRef90 bacterial clusters to UniRef50 clusters using DIAMOND blastp,⁴⁹ requiring >90% sequence identity and >90% coverage.

Statistical analysis and disease state prediction

Differentially abundant microbial features between patients and healthy controls were detected through generalised linear regression analysis using Maaslin2.⁵⁰ The linear model was formulated based on the log10-transformed abundances of each feature, with adjustments made for age and sex. The Benjamini-Hochberg method was employed to adjust for multiple testing. A Random Forest Classifier, predicting disease state from microbiome taxonomic and functional profiles, underwent a five-times repeated 10-fold cross-validation using scikit-learn.⁵¹ The classifier's hyperparameters were optimised via Randomized-SearchCV, using area under the receiver operating characteristic curve (AUROC) and area under the precision recall curve as the scoring metric, and tested across different studies. Prior to training, feature abundances underwent log10 transformation with small constant offsets ($1e^{-9}$) and were subsequently standardised using z-score normalisation. For individuals providing multiple samples from different visits, the sample from the first visit was selected for use in both intra-cross-validation and cross-studies validation.

Identification and analysis of human-microbiome PPIs

A previously established approach⁴⁰ was employed to map metagenomic datasets for predicting disease-relevant human-microbiome PPIs. For each individual in the study, bacterial proteins were identified and their abundances were aggregated based on their corresponding human protein interactors. Human proteins that were present in <5% of the study cohort were excluded. The identification of differentially abundant PPIs between patients and healthy controls was performed using generalised linear regression analysis using Maaslin2.⁵⁰

Human pathway annotation and enrichment analysis

Disease annotations were obtained from DisGeNET based on their gene-disease associations (GDAs), specifically those with GDA scores >0.1⁵² (June 2022). We performed pathway enrichment analysis using QIAGEN's Ingenuity Pathway Analysis software (IPA, QIAGEN Redwood City, California, USA, www.qiagen.com/ingenuity). Pathways were considered enriched if they had Benjamini-Hochberg-corrected p values <0.05.

Cell culture

HEK293T cells were obtained from American Type Culture Collection (CRL-3216) and maintained using Dulbecco's Modified Eagle Medium (Corning) supplemented with 10% fetal bovine serum (Cytiva). 18 hours before transfection, cells were seeded in a 24-well plate.

Plasmids

The FLAG-tagged expression vector pDEST-CMV-FLAG was generated by purchasing a vector from Addgene (#122845) and cloning out EGFP. The c-Myc-tagged expression vector pMH-MYC was purchased from Addgene (#101765). Human NR3C1

was Gateway cloned from pDONR-NR3C1 (CCSB Human Orfeome Collection) into pDEST-CMV-FLAG, human Hsp90 α was Gateway cloned from pDONR-Hsp90a (CCSB Human Orfeome Collection) into pMH-MYC, and all bacterial genes and mCherry were synthesised directly as Gene Fragments from Twist Bioscience and cloned into pDONR223, then Gateway cloned into pMH-MYC. All Gateway cloning was performed using Gateway LR Clonase II Enzyme mix (Invitrogen) according to the manufacturer's instructions. All plasmids were extracted using E.Z.N.A. Endo-free Plasmid DNA Mini Kit II (Omega Bio-Tek).

Co-immunoprecipitation assay

HEK293T cells in 24-well plate were transfected using JetPEI (Sartorius) according to the manufacturer's instructions. When co-transfecting pDEST-CMV-FLAG and pMH-MYC, only half the amount of recommended DNA was used for each vector. Media was replaced 12 hours after transfection. 48 hours post-transfection, cells were processed using Pierce c-Myc-Tag Magnetic IP/co-immunoprecipitation (Co-IP) Kit (Thermo Scientific) according to the manufacturer's instructions. Aliquots of the whole cell lysates and Co-IP elutions were flash frozen and stored at -80°C until further analysis.

Western blot analysis

Lysates and Co-IP elutions were subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis and transferred to nitrocellulose membranes using iBlot 2 Dry Blotting System (Invitrogen) for 7 min. Membranes were blocked using phosphate-buffered saline (PBS) with 5% non-fat milk and 0.1% Tween-20 for 1 hour then probed with an horseradish peroxidase (HRP)-conjugated anti-FLAG antibody (HRP-66008, ProteinTech) diluted 1:10 000 for 1.5 hours at room temperature. Membranes were washed three times with PBS+0.1% Tween-20. Proteins were visualised using Pierce ECL Western Blotting Substrate (Thermo Scientific). WesternSure Pre-stained Chemiluminescent Protein Ladder (LI-COR) was used as the molecular weight marker. Western blots were quantified using ImageJ.

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

RESULTS

Identification of taxonomic signatures associated with autoimmune disorders

To investigate whether an individual's gut microbiome composition is indicative of having an autoimmune disease, we analysed metagenomic datasets from seven studies focused on various autoimmune disorders (IBD),^{13 53} Graves' disease (GD),⁵⁴ SLE,³⁴ ankylosing spondylitis (AS)⁵⁵ and myasthenia gravis (MG),⁵⁶ in addition to one new SLE metagenomic cohort and one new MS metagenomic cohort⁴² (figure 1A; online supplemental table 1). The SLE cohort was the previous subject of longitudinal species-level analyses (with a focus on bacterial orthologs of the auto-antigen Ro60⁸). We included four published colorectal cancer (CRC) datasets for comparison, as microbiomes from individuals with CRC are typically dysbiotic.^{57 58} Additionally, since patients with IBD are at increased risk of developing CRC,⁵⁹ this predisposition may be reflected in their microbiomes. Combining datasets is often confounded by variables that are difficult to extract (sequencing depth, populational differences in diet, age, etc). Among the datasets we analysed, we noticed a strong bias in

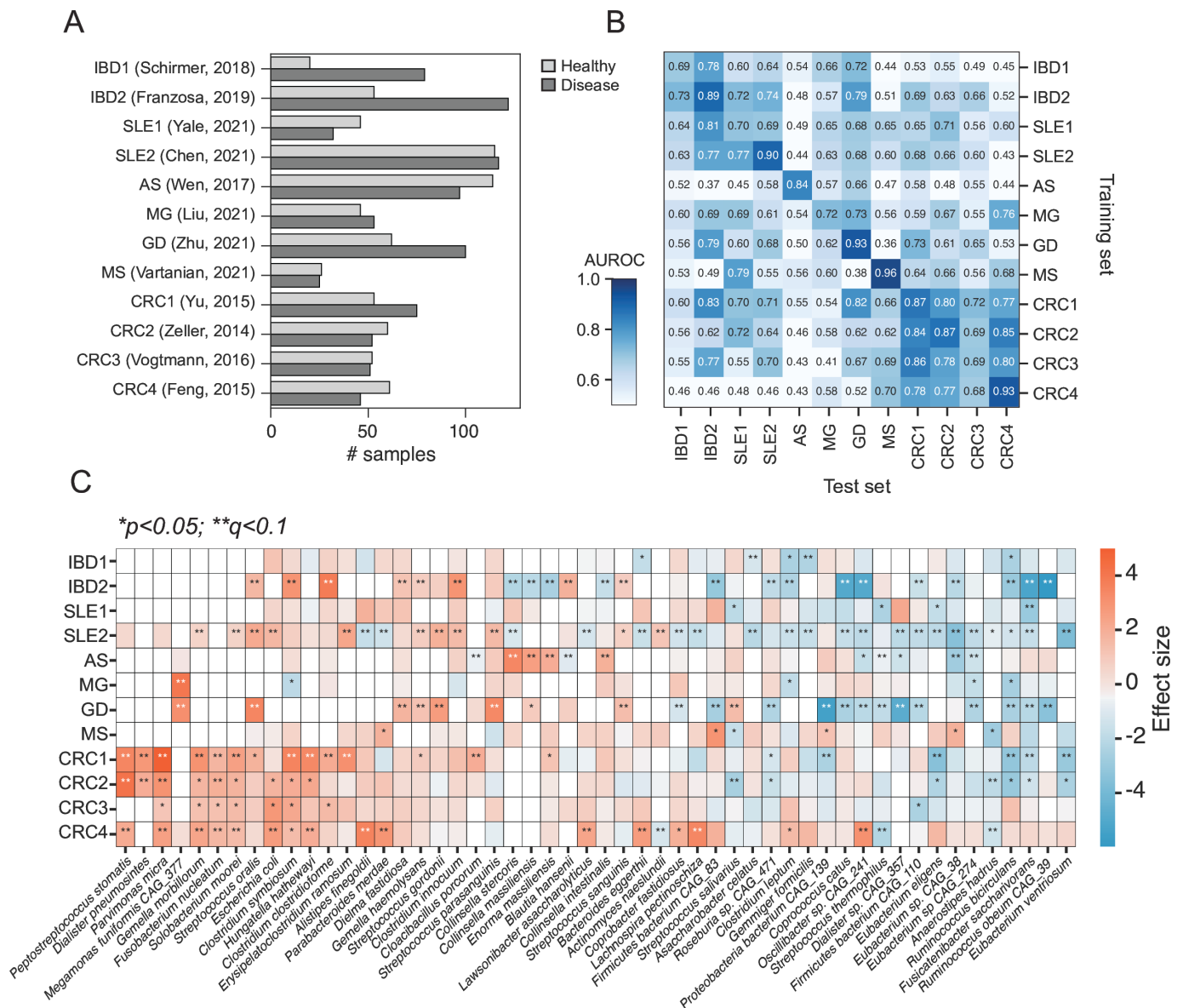


Figure 1 Taxonomic associations observed in metagenomic data from colorectal cancer (CRC) and autoimmune disease cohorts. (A) Each study charts the quantity of samples collected from both healthy and diseased individuals, with some individuals contributing multiple samples from various visits. (B) The area under the receiver operating characteristic curve (AUROC) for random forest models trained on the taxonomic composition from one cohort and used to predict labels (healthy/diseased) based on the taxonomic composition of individuals within the test cohort. For individuals providing multiple samples from different visits, the sample from the first visit was selected for use in both intra-cross-validation and cross-studies validation. Median values were calculated from multiple evaluations, encompassing both cross-study testing and fivefold cross-validation iterations. (C) Species with q values < 0.1 in two or more studies are plotted across all studies. Asterisks represent those studies in which the q value is < 0.1 . AS, ankylosing spondylitis; GD, Graves' disease; IBD, inflammatory bowel disease; MG, myasthenia gravis; MS, multiple sclerosis; SLE, systemic lupus erythematosus.

the phylum-level dominance of either Bacteroidota (formerly known as Bacteroidetes) or Bacillota (formerly known as Firmicutes) (online supplemental figure 1).

A useful method for identifying consistent features across datasets involves the utilisation of machine learning models in two distinct ways. First, by training models on one dataset and testing their predictive ability on another to test if the model generalises well beyond its training data. Second, we employ n -fold cross-validation to test the predictive power of the models within individual studies.^{60 61} We initially focused on the efficacy of random forest models trained on species compositions. In cross-cohort test-train scenarios, models trained with microbial compositions from patients with IBD could predict SLE

in other cohorts with a diagnostic ability ranging from modest to confident, as indicated by average AUROC values between 0.60 and 0.74. Conversely, using microbiome compositions of patients with SLE for training enabled the prediction of IBD with AUROCs between 0.70 and 0.90 (figure 1B, online supplemental figure 2). Interestingly, models trained on CRC datasets (CRC1; CRC2; CRC3; CRC4) demonstrated high predictive accuracy primarily within CRC datasets. Notably, intrastudy cross-validation revealed high performance across all datasets, with average AUROC scores ranging from 0.69 to 0.96.

Our analysis revealed common microbial species significant in both the IBD and SLE cohorts, especially pronounced in the IBD2 and SLE2 cohorts (figure 1C). We identified certain

species, like *Peptostreptococcus stomatis*, *Parvimonas micra*, *Gemella morbillorum*, *Fusobacterium nucleatum*, *Solobacterium moorei* and *Hungatella hathewayi*, as predominantly abundant in patients with CRC, some of which exacerbate CRC in mouse models,^{62 63} implying a possible pathogenic role. Our results highlighted species such as *Streptococcus oralis*, *Gemella haemolysans* and *Clostridium innocuum*, which were more abundant in patients with IBD and SLE compared with controls (online supplemental figure 3; online supplemental table 2). In contrast, control groups in IBD and SLE studies exhibited higher abundances of certain microbial species, indicating a baseline microbial community composition in healthy individuals that is altered in autoimmune diseases. *Anaerostipes hadrus* was one of few species consistently more abundant in healthy individuals compared with those with CRC (false discovery rate (FDR) <0.1). Importantly, *Fusicatenibacter saccharivorans* was more abundant in healthy individuals compared with patients with multiple diseases including IBD, SLE, GD and CRC.

Furthermore, we observed a higher abundance of species such as *Eubacterium* sp CAG_38, *Gemmiger formicilis*, *C. leptum* and *Asaccharobacter celatus* in healthy controls compared with patients with IBD and SLE. This finding aligns with reduced abundance of *C. leptum* in the faecal microbiota of patients with UC compared with healthy individuals.⁶⁴ In addition, other studies have suggested that supplementation with butyrate, which is produced by *C. leptum* and other *Clostridium* cluster XIVa bacteria, may have therapeutic benefits in the treatment of IBD and other autoimmune diseases.⁶⁵ However, while these models elucidate taxonomic-level commonalities between SLE and IBD, they fall short of explaining the underlying mechanisms of these relationships.

Microbial functions predict autoimmune disease

Microbial functions that have been thus far shown to exert influence on the immune system and alter the pathogenesis of autoimmune disease include metabolism, production of bacteriocins, modification of host structures/enzymes, gene regulation, competition between microbiota and host for nutrients and/or space and the synthesis of secondary metabolites (eg, organic acids, bile acids).^{66 67} To predict microbial gene families that associate specifically with autoimmune diseases, we first used models trained on the composition of protein families (PFAMs) in each cohort's microbiomes (figure 2A). PFAMs provide functional insight into the mechanisms that may be involved in specific disease states. We identified key PFAMs significantly more abundant in healthy controls in studies focusing on SLE and IBD, notably 'PF00404; Dockerin type I domain', 'PF12891; Glycoside hydrolase family 44' and 'PF08672; Anaphase promoting complex subunit 2' (online supplemental table 3). Both the type I dockerin domain and glycoside hydrolase family 44 proteins are important in the degradation of cellulose, suggesting a role for fibre degradation in maintaining intestinal health.

To examine fibre degradation more closely, we next specifically analysed carbohydrate-active enzyme (CAZymes) profiles (figure 2B; online supplemental table 4). We identified several CAZymes (ie, lipopolysaccharide N-acetylglucosaminyltransferase (GT9), peptidoglycan hydrolase (GH73)) that were significantly enriched in patients with multiple autoimmune diseases and CRC (online supplemental figure 4). This enrichment suggests a potential link between these CAZymes and disease pathogenesis, possibly through modulation of gut microbiota composition and function.^{68 69} Additionally, genes encoding enzymes such as endo- α -1,4-polygalactosaminidase

(GH114), endo- β -1,4-galactanase (GH53) and glucan endo-1,3- β -glucosidase (GH17) were consistently more abundant in healthy controls within SLE and IBD datasets,⁷⁰ highlighting the potential for CAZymes to be used as biomarkers for the diagnosis of autoimmune conditions like IBD and SLE.

Many species of gut bacteria produce SCFAs that act as anti-inflammatory signals for immune cells. Pyruvate is anti-inflammatory and is commonly used to treat inflammatory conditions.⁷¹ Studies have shown that ethyl pyruvate, a derivative of pyruvate, can ameliorate various forms of inflammatory liver injury, suggesting its role in reducing systemic inflammation caused by the liver.⁷² The role of butyrate in promoting regulatory T cell functions contributes to the maintenance of gut homeostasis and the suppression of inflammation.⁷³ Butyrate also influences the activity of cytotoxic T lymphocytes (CTLs). Contrary to enhancing the production of effector cytokines, butyrate has been shown to modulate CTL activity in a manner that reduces inflammatory responses, aligning with its broader anti-inflammatory effects.⁷⁴ Furthermore, butyrate has been found to inhibit the activation of antigen-specific CD8⁺ T cells by affecting the antigen-presenting dendritic cells, leading to a reduction in systemic inflammation.⁷⁵ Our analysis revealed significant variances in the abundance of SCFA-producing microbial enzymes between control groups and patients with SLE and IBD. We identified 66 enzymes related to SCFA production, with 33 being common between the two diseases, showing higher levels in patient groups. Conversely, 18 enzymes, with 8 common between SLE and IBD, were more prevalent in healthy controls. This pattern indicates a potential involvement of these enzymes in the pathogenesis of SLE and IBD (online supplemental table 5).

Our analysis uncovers a marked metabolic distinction between healthy individuals and patients with SLE and IBD, particularly in pyruvate and acetyl-CoA metabolism.^{76 77} Patients with SLE and IBD show an enrichment of enzymes such as pyruvate dehydrogenase and pyruvate kinase, shifting the metabolic focus towards SCFA production from pyruvate, which may influence disease progression through gut microbiota alterations. Similarly, acetyl-CoA metabolism in healthy controls is characterised by a dominance of enzymes like citrate (Re)-synthase, supporting the TCA cycle, whereas patients exhibit higher levels of enzymes like acetate CoA-transferase, diverting acetyl-CoA towards acetate production, an SCFA, and potentially affecting microbial composition and inflammation.⁷⁸ Furthermore, the analysis indicates specific increases in succinate dehydrogenase and propionyl-CoA synthetase in patients, suggesting a reoriented metabolism towards succinate and propionate pathways, implicating these SCFAs in modulating immune responses and inflammatory conditions.⁷⁹

Importantly, since functions can be redundant across organisms, they reflect an orthogonal axis for analysis. We compared their relative utility as training data for classifying each disease, using AUROC as our primary metric (figure 2C). Whereas taxonomic profiles are more successful at classifying CRC status, gene functions (PFAMs) provided significantly more discriminatory value in classifying autoimmune diseases, including SLE, GD, MG and AS.

Host-microbiome PPIs associated with autoimmune diseases capture disease pathways

We first identify human proteins that bind to microbiome proteins in each cohort based on their homology to proteins that have been experimentally shown to bind these human proteins.

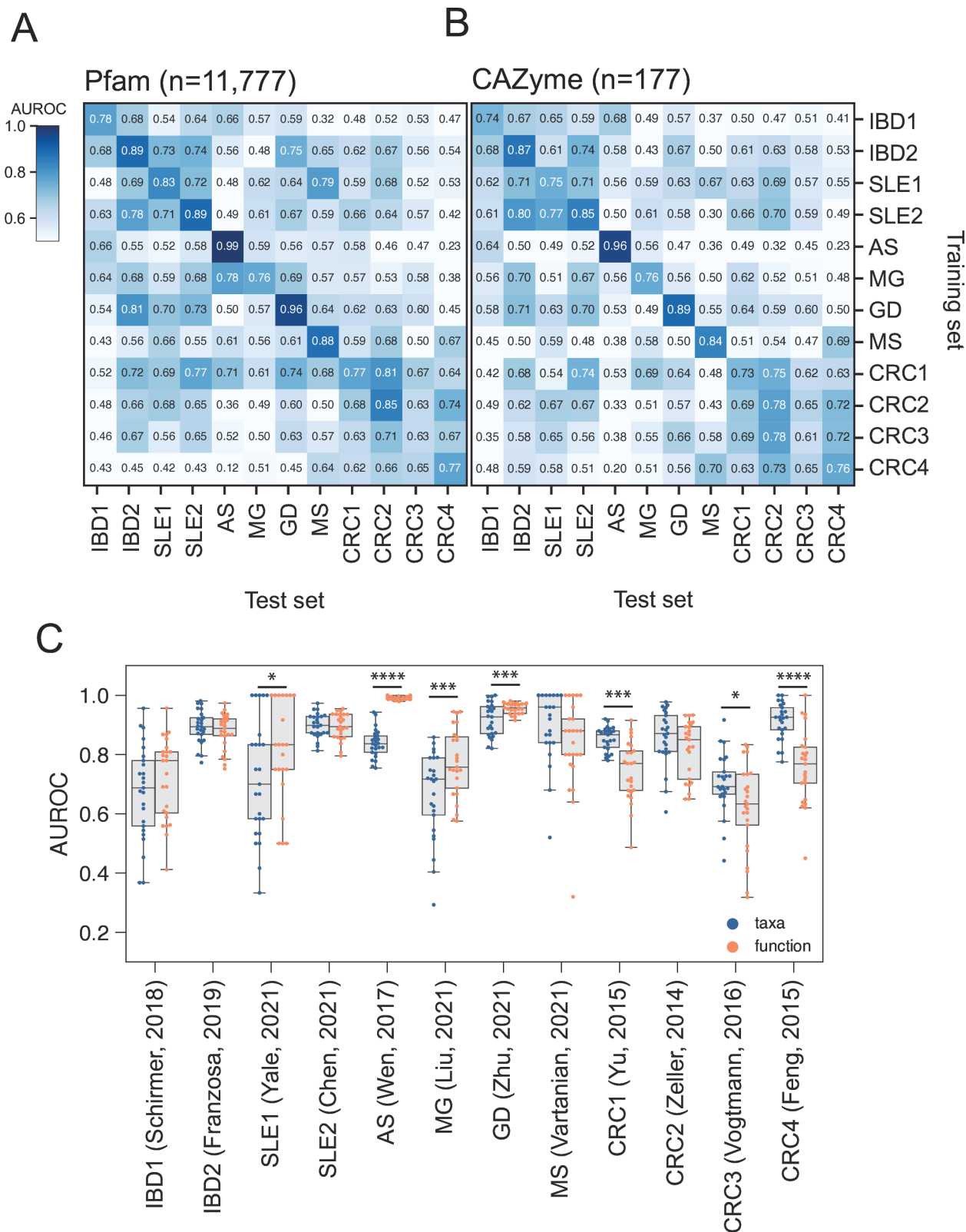


Figure 2 Microbial functional genes and CAZymes exhibit comparable predictive performance to species for autoimmune diseases. (A) The area under the receiver operating characteristic curve (AUROC) for random forest models trained on the protein family abundances, determined using Pfam, from one cohort and used to predict labels (healthy/diseased) based on the protein family abundances of individuals within the test cohort. (B) AUROC for random forest models trained on the CAZyme profiles from one cohort and used to predict labels (healthy/diseased) based on the CAZyme profiles of individuals within the test cohort. (C) Box plot comparison of predictive performance when taxonomic or functional abundance profiles are used in microbiome datasets. The AUROC was estimated through a fivefold cross-validation experiment, conducted five times. Welch's t-test was used to determine the statistical significance between the two approaches. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. AS, ankylosing spondylitis; CRC, colorectal cancer; GD, Graves' disease; IBD, inflammatory bowel disease; MG, myasthenia gravis; MS, multiple sclerosis; SLE, systemic lupus erythematosus.

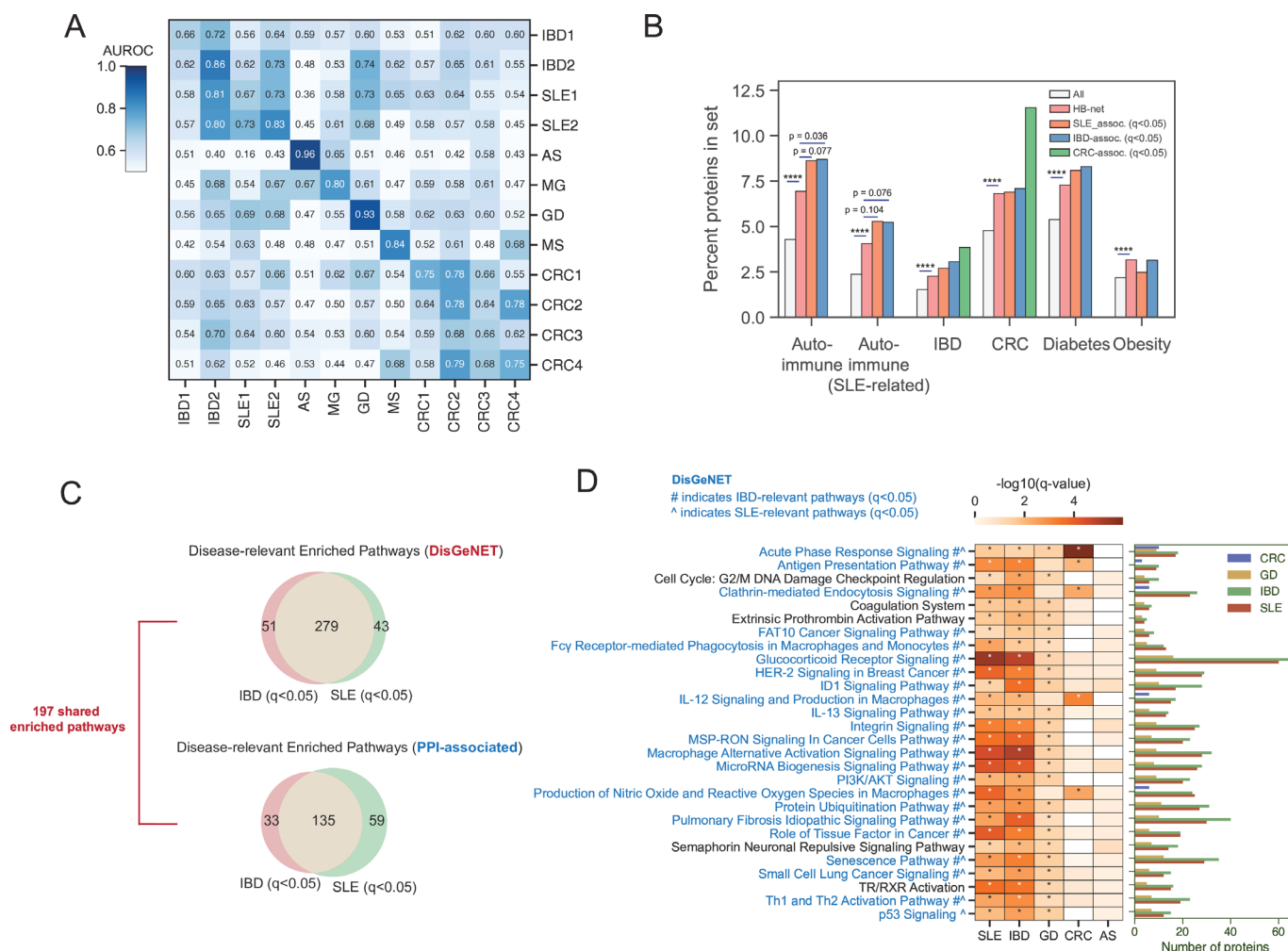


Figure 3 Host-microbiome PPIs provide insight into disease processes associated with IBD and SLE. (A) Area under the receiver operating characteristic curve (AUROC) for random forest models trained on the summed abundance of bacterial interactors which target each human protein interactor from one cohort and used to predict labels (healthy/diseased) based on the same information from individuals within the test cohort. (B) Comparison of PPI-associated genes and the percentage of these genes identified as disease-associated in DisGeNET, in relation to the predicted host-microbiome PPIs (HB-net) and all reviewed human proteins from Uniprot. (C) Venn diagrams showing the overlap of enriched pathways relevant to IBD and SLE, based on genes associated with these diseases according to DisGeNET or disease-associated PPIs ($q < 0.05$). (D) Enriched pathways associated with human protein interactors identified as important features in each disease type are plotted. Only those pathways associated with three or more diseases are plotted in the heatmap according to their Benjamini-Hochberg-adjusted p values. Asterisks indicate Benjamini-Hochberg-adjusted p values < 0.05 . Pathways known to play a role in IBD and SLE are marked in red and labelled by a hashtag and/or asterisk, respectively. The total number of human proteins identified as important features within each pathway are plotted according to the disease. AS, ankylosing spondylitis; CRC, colorectal cancer; GD, Graves' disease; IBD, inflammatory bowel disease; MG, myasthenia gravis; MS, multiple sclerosis; PPI, protein-protein interaction; SLE, systemic lupus erythematosus.

Using the same methodology as we applied to train models on individual cohorts and test their ability to classify disease status of other cohorts, we recapitulated our previous findings, namely that there is strong predictive power across CRC cohorts (AUROC 0.55–0.79) and between IBD and SLE cohorts (AUROC 0.56–0.86) (figure 3A). We identified human proteins with known gene disease-associated mutations using DisGeNET.⁸⁰ For the microbiome-associated disorders that we analysed, we found an enrichment of genes with GDAs. We found a modestly significant enrichment of PPIs involving human proteins with GDAs with autoimmune disorders in the SLE-associated and IBD-associated metagenomes (figure 3B). Overall, there were at least three times more human targets, which implicate specific disease pathways,⁴⁰ identified in the IBD and SLE cohorts than any of the other diseases (figure 3C), which further suggests

that these diseases are driven by functional components across different species.

The identification of recurring modules associated with autoimmune diseases through the analysis of host-microbiome PPIs provides additional evidence for specific microbial mechanisms involved in disease pathogenesis. We annotated pathways for the human protein interactors using Ingenuity Pathway Analysis.⁸¹ Various immune response pathways, such as glucocorticoid receptor signalling, interleukin (IL)-12 signalling, IL-13 signalling and PI3K/AKT signalling, were found to be enriched in both SLE and IBD (figure 3D), suggesting common underlying pathophysiologies.⁸² On the other hand, we did not find any significant host-microbiome PPIs associated with AS. These results suggest that host-microbiome PPIs may be more specific to each autoimmune disease, rather than universally associated

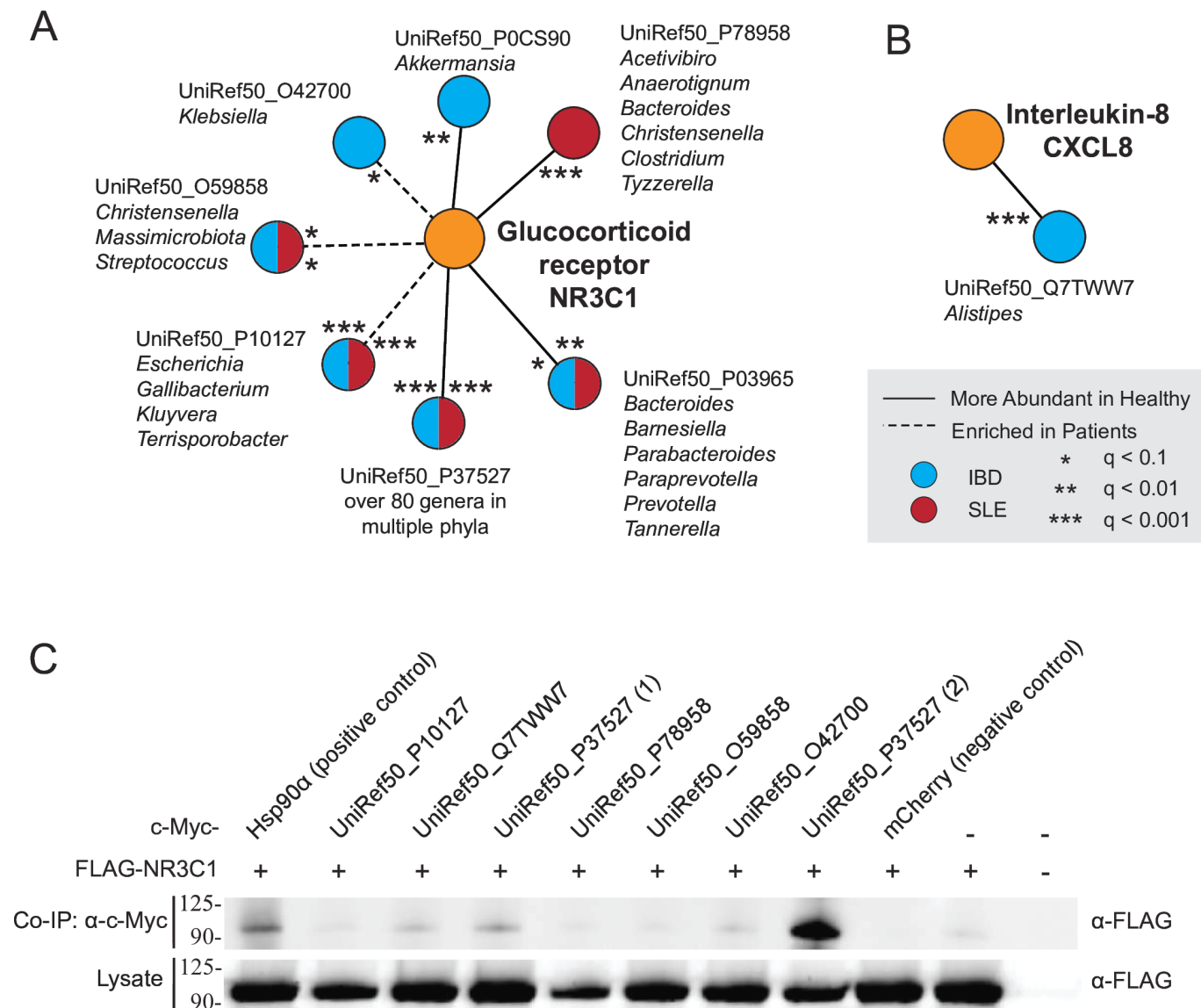


Figure 4 Predicted human-microbiome PPIs linked to IBD and SLE. (A) Mapping of NR3C1-associated human-microbiome PPIs relevant to disease. Displayed are only those PPIs with significant associations and concurrent bacterial protein clusters linked to disease. The q values represent p values adjusted using the Benjamini-Hochberg (BH) method. The metagenomics datasets' genera predicted to contain these UniProt clusters are annotated. (B) Mapping of CXCL8 (C-X-C motif chemokine ligand 8)-associated human-microbiome PPIs pertinent to disease. (C) HEK293T cells expressing FLAG-NR3C1 and c-Myc-tagged controls/bacterial proteins, FLAG-NR3C1 alone or nothing were subject to co-immunoprecipitation (Co-IP) using an anti-c-Myc Co-IP kit. Whole cell lysates and the Co-IP elutions were subject to western blot analysis using an anti-FLAG antibody to identify FLAG-NR3C1. (1) and (2) indicate the first and second representatives from the same UniRef50 cluster.

with autoimmunity. Further research is necessary to determine if these observations can be replicated in larger and more diverse cohorts.

Next, we focused our analysis on specific interactions that were common between SLE and IBD. Glucocorticoid receptor (NR3C1 in UniProt) is essential as a therapeutic target for both conditions, serving as the primary receptor for glucocorticoid hormones and analogues—often employed to suppress chronic inflammation in IBD and SLE. The interaction between glucocorticoid hormones and the NR3C1 receptor aids in regulating immune response and inflammation, thereby managing disease progression. Fang *et al.*⁸³ highlight the significance of NR3C1 in the context of autoimmune disorders, emphasising its role in glucocorticoid resistance, a concern in the treatment of chronic inflammatory diseases.⁸³ Predicted host-microbiome PPIs relevant to NR3C1 were significantly associated in both IBD and

SLE studies (SLE2: FDR < 1e⁻⁴; IBD2: FDR < 0.01; figure 4A). We identified potential PPIs between NR3C1 and a bacterial cluster encoding glutathione peroxidase-like peroxiredoxin Gpx1 (UniRef50_O59858), which was significantly enriched in patients with IBD and SLE (SLE2: FDR < 0.1; IBD2: FDR < 0.05). This interaction could potentially modulate the oxidative stress response, thereby influencing inflammation and tissue damage in these conditions. Another interaction between NR3C1 and a bacterial cluster encoding pyridoxal 5'-phosphate synthase subunit PdxS (UniRef50_P37527) was significantly more abundant in healthy controls in IBD and SLE studies (SLE2: FDR < 10⁻⁴; IBD2: FDR < 10⁻³). This PPI may be important for vitamin B₆ metabolism, which has been implicated in immune function and inflammation,⁸⁴ suggesting a potential role of the microbiome in modulating disease mechanisms and therapeutic responses. All significant interactions (FDR < 0.1) between

bacterial proteins and NR3C1 were identified in the SLE2 and IBD2 cohorts. Interestingly, the direction of the enrichment for each bacterial protein-NR3C1 interaction was consistent between both SLE2 and IBD2 studies (online supplemental table 6).

The diversity of species associated with PPIs range from those common across various lineages to those specific to a narrow lineage. This variability is exemplified in the interaction of bacterial proteins with NR3C1, a nuclear receptor. The protein UniRef50_P78958, glyceraldehyde-3-phosphate dehydrogenase 1, for example, is identified in a broad array of genera such as *Acetivibrio*, *Anaerotignum*, *Bacteroides*, *Christensenella*, *Clostridium* and *Tyzzzeria*. Likewise, UniRef50_P37527, the pyridoxal 5'-phosphate synthase subunit PdxS, showcases an even broader distribution, with its presence identified in over 80 different genera across multiple phyla, reflecting a highly diverse range of microbial interactions.

IL-8, also known as CXCL8, is implicated in the inflammatory processes of both IBD and SLE. In IBD, it directs neutrophil migration to the inflammation site in the gut, leading to persistent inflammation and tissue damage.⁸⁵ Similarly, in SLE, elevated CXCL8 levels stimulate neutrophil migration to tissues, resulting in the characteristic inflammation and tissue damage seen in these patients. A meta-analysis by Mao *et al*⁸⁶ confirms the increased circulating levels of CXCL8 in patients with SLE, reinforcing its role in the disease's pathophysiology.^{86 87} For CXCL8, significant host-microbiome PPIs were similarly observed in both IBD and SLE studies (SLE2: FDR <0.05; IBD2: FDR <0.001; IBD1: $p < 0.05$, figure 4B). However, the scope of these interactions was more limited. The most notable was the association of the UniRef50_Q7TWW7 cluster, identified as adenosylhomocysteinase encoded by *Alistipes* sp, with IBD. This cluster showed significant enrichment in healthy controls (IBD1: $p < 0.01$; IBD2: FDR <10⁻⁴). Whether this enzyme is involved in the regulation of homocysteine levels—known to be associated with inflammation and autoimmune diseases⁸⁸—or in binding a protein known to play a role in autoimmune disease remains to be determined experimentally.

To confirm binding between NR3C1 and the predicted bacterial interactors in vivo, we expressed FLAG-tagged NR3C1 and c-Myc-tagged bacterial proteins in HEK293T cells and performed a co-immunoprecipitation assay (figure 4C; online supplemental figure 5). For each tested UniRef50 cluster (online supplemental table 6), we selected 1–2 representative protein sequences from bacterial species detected in the gut microbiomes of patients with IBD and SLE for further testing. FLAG-NR3C1 clearly co-immunoprecipitated with both representatives of UniRef50_P37527 (PdxS; A0A0M6WFX4; *Roseburia faecis* and, to a lesser extent, PdxS; U2KEI3; *Ruminococcus callidus*). Additionally, adenosylhomocysteinase (B0MW70; *Alistipes putredinis*) of UniRef50_Q7TWW7 and transaldolase (A0A2X4Z5B2; *Klebsiella oxytoca*) of UniRef50_O42700 showed binding to FLAG-NR3C1, although the signal was notably weaker than that of the positive control Hsp90a, a known interactor with NR3C1.⁸⁹ Interestingly, proteins GapA/N1ZEC1 from UniRef50_P78958 and PdxS/U2KEI3 from UniRef50_P37527 exhibited smaller-sized bands, suggesting potential proteolytic degradation of NR3C1 (online supplemental figure 6). These findings suggest that NR3C1 binds with predicted microbial interactors. Further experiments are warranted to determine the functional implications of these interactions.

DISCUSSION

This study offers critical insights into the role of the gut microbiota in autoimmune diseases. By using machine learning models

and analysing taxonomic signatures, microbial functions and host-microbiome PPIs in patients with various autoimmune diseases, we identified shared microbial features and functional contributions between SLE and IBD. Although SLE and IBD share some clinical aspects and treatment regimens, the co-occurrence of these two diseases is uncommon.⁹⁰ Furthermore, whereas UC is generally restricted to the colon, the small intestinal microbiota is implicated in the pathogenesis of SLE.²³ Specifically, translocation of pathobionts causally implicated in SLE translocate from the small intestine to extraintestinal sites in the body.^{21 23 91} As our analysis focuses primarily on the stool microbiome, a proxy for large intestinal communities, we posit that some functions are shared between the small and large intestines in patients with SLE or that residual signal from small intestinal microbiota can be found within the stool.

In addition to the similarities observed between PPIs in SLE and IBD patient cohorts, prior work on B cell receptor repertoires has revealed similar increases in B cell clones in both SLE and CD, dominated by the IgA isotype, particularly the IgA1 and IgA2 subclasses.⁹² This suggests a mucosa-derived microbial contribution to the pathogenesis of both disorders. Importantly, a skewed immunoglobulin heavy chain variable region (IGHV) gene usage was observed in both disease states, in particular IGHV4-34 and VH4-59, which are both autoreactive, with IGHV4-34 having been demonstrated to also bind to gut commensal bacteria.⁹³ In conjunction with our results on shared PPIs in both SLE and IBD, one could speculate that these interactions may contribute to the induction of similar adaptive immune responses. Further studies are needed to test whether shared adaptive immune responses (such as the IL-12-related T helper type 1 differentiation pathway or particular B cell clones) are gut microbiota-driven. On the innate immune side, subsets of patients with SLE and IBD share a type I interferon signature, which is particularly dominant in SLE and may also be aggravated by gut pathobionts beyond known host genetic contributions.^{21 94–96}

Furthermore, our findings suggest several coherent concepts with potential therapeutic implications. First, certain fibre-degrading enzymes were consistently more abundant in healthy individuals across multiple studies, highlighting the importance of dietary fibre in maintaining gut health. This suggests that high-fibre diets or supplementation with fibre-degrading probiotics could help restore gut homeostasis and potentially mitigate autoimmune symptoms.^{97 98}

Second, the analysis revealed significant differences in the abundance of SCFA-producing microbial enzymes between healthy controls and autoimmune patients. SCFAs, such as butyrate, play crucial roles in regulating immune responses and maintaining intestinal health. Therapeutic approaches could include SCFA supplementation or promoting the growth of SCFA-producing bacteria through prebiotics to reduce inflammation and support immune regulation in autoimmune diseases.^{99 100}

Third, the identification of key host-microbiome PPIs, particularly those involving NR3C1, reaffirms previously known host factors and introduces interactions identified through our approach that may offer additional therapeutic opportunities. These interactions are likely to occur in immune cells, such as macrophages and lymphocytes, where NR3C1 plays a critical role in regulating immune responses.¹⁰¹ The effect of these bacterial proteins on NR3C1-mediated gene transcription could influence a wide range of biological processes, including glucocorticoid signalling, which impacts immune cell function, inflammation and stress responses. Given NR3C1's role as a multitasking transcription factor, these interactions may alter the balance between

its anti-inflammatory and pro-inflammatory actions, potentially contributing to the pathogenesis or treatment of autoimmune diseases. For example, microbial proteins that interact with NR3C1 and modulate oxidative stress responses and vitamin B₆ metabolism could be targeted to enhance glucocorticoid therapy efficacy.

Additionally, targeting bacterial proteins that might bind to CXCL8 could be explored as a way to disrupt pro-inflammatory pathways while potentially preserving CXCL8's role in neutrophil recruitment and immune response.^{5,87} Importantly, beyond potential future therapeutic opportunities, the effect of the human gut microbiota on otherwise well-studied host pathways such as those related to corticosteroids is little known and understood. Few mechanistic studies support effects of the microbiota on glucocorticoid metabolism and functions.^{102,103} The findings of our study imply that additional host effects from the gut microbiota may occur, although further mechanistic work is needed to solidify PPIs among bacteria and this important host pathway known to impact inflammatory diseases.

In conclusion, our study provides strong evidence for shared microbial signatures and functional alterations across different autoimmune diseases, highlighting the microbiome as a potential therapeutic target. We have identified specific microbial features, functional pathways and host-microbiome PPIs that warrant further investigation. This research lays a foundation for developing novel microbiome-based interventions, ranging from dietary modifications to targeted probiotics, prebiotics, faecal microbiota transplantation and host-microbiome PPI modulation, offering new hope for the prevention and treatment of a broad range of autoimmune diseases.

Data and code availability statement

The sequencing data has been deposited in the NCBI database under the accession number PRJNA1156897. Additionally, the code employed for training predictive models can be found at: https://github.com/britolab/ppi_autoimmune.¹⁰⁴

Acknowledgements The authors would like to thank Márcia Pereira for insightful comments on this manuscript, Teri Greiling and Carina Dehner for initial SLE cohort development and Iryna Kulyk for SLEDAI calculations.

Contributors HZ, MAK and IB designed the study. TV and MAK oversaw patient samples collection. DB performed the co-immunoprecipitation experiment. QS and DB performed the metagenomic sequencing. HZ and IB designed the analysis plan and performed computational analysis. All authors revised the manuscript. All authors contributed to the article and approved the submitted version. IB is the guarantor.

Funding This work was primarily funded by the Lupus Research Alliance-Bristol-Myers Squibb Accelerator award. IB was also supported by grants from the National Institutes of Health (1DP2HL141007), and the Lupus Research Alliance Global Team Science Award. MAK was partly supported by grants from the Lupus Research Alliance (Lupus Insight Prize, Global Team Science Award), the National Institutes of Health (NIH R01AI118855), Arthritis National Research Foundation, Arthritis Foundation and the Maren Foundation.

Competing interests MAK received salary, consulting fees, honoraria or research funds from Eligo Biosciences, Enterome, Novartis, Roche, Genentech, Bristol-Meyers Squibb, Sanofi, and AbbVie, and holds a patent on the use of microbiota manipulations to treat immune-mediated diseases. HZ is both a salaried employee and a shareholder of Moderna.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval In accordance with the Yale Human Investigations Committee (#1408014402) and Cornell University Institutional Review Board (#2006009658), fresh stool samples were gathered from participants who were fully informed and gave their consent.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iD

Ilana Brito <http://orcid.org/0000-0002-2250-3480>

REFERENCES

- Halfvarson J, Brislawn CJ, Lamendella R, et al. Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol* 2017;2:17004.
- Zhang X, Zhang D, Jia H, et al. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *N Med* 2015;21:895–905.
- Ruff WE, Greiling TM, Krieger MA. Host-microbiota interactions in immune-mediated diseases. *Nat Rev Microbiol* 2020;18:521–38.
- Pan Q, Guo F, Huang Y, et al. Gut Microbiota Dysbiosis in Systemic Lupus Erythematosus: Novel Insights into Mechanisms and Promising Therapeutic Strategies. *Front Immunol* 2021;12:799788.
- Miyauchi E, Shimokawa C, Steimle A, et al. The impact of the gut microbiome on extra-intestinal autoimmune diseases. *Nat Rev Immunol* 2023;23:9–23.
- Tomofuji Y, Maeda Y, Oguro-Igashira E, et al. Metagenome-wide association study revealed disease-specific landscape of the gut microbiome of systemic lupus erythematosus in Japanese. *Ann Rheum Dis* 2021;80:1575–83.
- Schirmer M, Franzosa EA, Lloyd-Price J, et al. Dynamics of metatranscription in the inflammatory bowel disease gut microbiome. *Nat Microbiol* 2018;3:337–46.
- Greiling TM, Dehner C, Chen X, et al. Commensal orthologs of the human autoantigen Ro60 as triggers of autoimmunity in lupus. *Sci Transl Med* 2018;10:eaan2306.
- Gronke K, Nguyen M, Santamaria N, et al. Human Th17- and IgG3-associated autoimmunity induced by a translocating gut pathobiont. *bioRxiv* 2023.:2023.06.29.546430.
- Longman RS, Littman DR. The functional impact of the intestinal microbiome on mucosal immunity and systemic autoimmunity. *Curr Opin Rheumatol* 2015;27:381–7.
- Chen B, Du G, Guo J, et al. Bugs, drugs, and cancer: can the microbiome be a potential therapeutic target for cancer management? *Drug Discov Today* 2019;24:1000–9.
- Ryu TY, Kim K, Han T-S, et al. Human gut-microbiome-derived propionate coordinates proteasomal degradation via HECTD2 upregulation to target EHMT2 in colorectal cancer. *ISME J* 2022;16:1205–21.
- Franzosa EA, Sirota-Madi A, Avila-Pacheco J, et al. Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol* 2019;4:293–305.
- De Luca F, Shoenfeld Y. The microbiome in autoimmune diseases. *Clin Exp Immunol* 2019;195:74–85.
- Pereira MS, Krieger MA. Evolving concepts of host-pathobiont interactions in autoimmunity. *Curr Opin Immunol* 2023;80:102265.
- Hevia A, Milani C, López P, et al. Intestinal dysbiosis associated with systemic lupus erythematosus. *MBio* 2014;5:e01548–14.
- Azzouz DF, Chen Z, Izmirly PM, et al. Longitudinal gut microbiome analyses and blooms of pathogenic strains during lupus disease flares. *Ann Rheum Dis* 2023;82:1315–27.
- Chen B, Sun L, Zhang X. Integration of microbiome and epigenome to decipher the pathogenesis of autoimmune diseases. *J Autoimmun* 2017;83:31–42.
- Rogers GB. Germs and joints: the contribution of the human microbiome to rheumatoid arthritis. *N Med* 2015;21:839–41.
- Dehner C, Fine R, Krieger MA. The microbiome in systemic autoimmune disease: mechanistic insights from recent studies. *Curr Opin Rheumatol* 2019;31:201–7.
- Zegarra-Ruiz DF, El Beidaq A, Iñiguez AJ, et al. A Diet-Sensitive Commensal Lactobacillus Strain Mediates TLR7-Dependent Systemic Autoimmunity. *Cell Host Microbe* 2019;25:113–27.
- Nogueira AR, Shoenfeld Y. Microbiome and autoimmune diseases: cause and effect relationship. *Curr Opin Rheumatol* 2019;31:471–4.

- 23 Manfredo Vieira S, Hiltensperger M, Kumar V, *et al.* Translocation of a gut pathobiont drives autoimmunity in mice and humans. *Science* 2018;359:1156–61.
- 24 Bagavant H, Araszewicz AM, Ingram JK, *et al.* Immune Response to *Enterococcus gallinarum* in Lupus Patients Is Associated With a Subset of Lupus-Associated Autoantibodies. *Front Immunol* 2021;12:635072.
- 25 Nakamoto N, Sasaki N, Aoki R, *et al.* Gut pathobionts underlie intestinal barrier dysfunction and liver T helper 17 cell immune response in primary sclerosing cholangitis. *Nat Microbiol* 2019;4:492–503.
- 26 Azzouz D, Omarbekova A, Heguy A, *et al.* Lupus nephritis is linked to disease-activity associated expansions and immunity to a gut commensal. *Ann Rheum Dis* 2019;78:947–56.
- 27 Ishaq HM, Mohammad IS, Shahzad M, *et al.* Molecular Alteration Analysis of Human Gut Microbial Composition in Graves' disease Patients. *Int J Biol Sci* 2018;14:1558–70.
- 28 Mu Q, Zhang H, Liao X, *et al.* Control of lupus nephritis by changes of gut microbiota. *Microbiome* 2017;5:73.
- 29 Schnabl B, Brenner DA. Interactions between the intestinal microbiome and liver diseases. *Gastroenterology* 2014;146:1513–24.
- 30 Duvallet C, Gibbons SM, Gurry T, *et al.* Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat Commun* 2017;8:1784.
- 31 Volkova A, Ruggles KV. Predictive Metagenomic Analysis of Autoimmune Disease Identifies Robust Autoimmunity and Disease Specific Microbial Signatures. *Front Microbiol* 2021;12:621310.
- 32 Ning L, Zhou Y-L, Sun H, *et al.* Microbiome and metabolome features in inflammatory bowel disease via multi-omics integration analyses across cohorts. *Nat Commun* 2023;14:7135.
- 33 Mills RH, Dulai PS, Vázquez-Baeza Y, *et al.* Multi-omics analyses of the ulcerative colitis gut microbiome link *Bacteroides vulgatus* proteases with disease severity. *Nat Microbiol* 2022;7:262–76.
- 34 Chen B, Jia X, Xu J, *et al.* An Autoimmunogenic and Proinflammatory Profile Defined by the Gut Microbiota of Patients With Untreated Systemic Lupus Erythematosus. *Arthritis Rheumatol* 2021;73:232–43.
- 35 Wen S, Peng S, Hu X, *et al.* Validation of metagenomic next-generation sequencing of bronchoalveolar lavage fluid for diagnosis of suspected pulmonary infections in patients with systemic autoimmune rheumatic diseases receiving immunosuppressant therapy. *Front Med (Lausanne)* 2023;10:1161661.
- 36 König MF, Abusleme L, Reinholdt J, *et al.* Aggregatibacter actinomycetemcomitans-induced hypercitrullination links periodontal infection to autoimmunity in rheumatoid arthritis. *Sci Transl Med* 2016;8:369ra176.
- 37 Obata T, Goto Y, Kunisawa J, *et al.* Indigenous opportunistic bacteria inhabit mammalian gut-associated lymphoid tissues and share a mucosal antibody-mediated symbiosis. *Proc Natl Acad Sci U S A* 2010;107:7419–24.
- 38 Thomas CM, Versalovic J. Probiotics-host communication. *Gut Microbes* 2010;1:148–63.
- 39 Hwang IY, Chang MW. Engineering commensal bacteria to rewire host-microbiome interactions. *Curr Opin Biotechnol* 2020;62:116–22.
- 40 Zhou H, Beltrán JF, Brito IL. Host-microbiome protein-protein interactions capture disease-relevant pathways. *Genome Biol* 2022;23:72.
- 41 Petri M, Kim MY, Kalunian KC, *et al.* Combined oral contraceptives in women with systemic lupus erythematosus. *N Engl J Med* 2005;353:2550–8.
- 42 Ma Y, Sannino D, Linden JR, *et al.* Epsilon toxin-producing *Clostridium perfringens* colonize the multiple sclerosis gut microbiome overcoming CNS immune privilege. *J Clin Invest* 2023;133:e163239.
- 43 Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 2011;27:863–4.
- 44 BMTagger. Available: <https://ftp.ncbi.nlm.nih.gov/pub/agarwala/bmtagger/>
- 45 Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–20.
- 46 Beghini F, McIver LJ, Blanco-Míguez A, *et al.* Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *Elife* 2021;10:e65088.
- 47 Zhang J, Zuo K, Fang C, *et al.* Altered synthesis of genes associated with short-chain fatty acids in the gut of patients with atrial fibrillation. *BMC Genomics* 2021;22:634.
- 48 Pasolli E, Schiffer L, Manghi P, *et al.* Accessible, curated metagenomic data through ExperimentHub. *Nat Methods* 2017;14:1023–4.
- 49 Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 2015;12:59–60.
- 50 Mallick H, Rahnavard A, McIver LJ, *et al.* Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput Biol* 2021;17:e1009442.
- 51 Pedregosa F, Varoquaux G, Gramfort A, *et al.* Scikit-learn: machine learning in python. *arXiv* 2018. Available: <http://arxiv.org/abs/1201.0490>
- 52 Piñero J, Bravo A, Queralt-Rosinach N, *et al.* DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res* 2017;45:D833–9.
- 53 Schirmer M, Denson L, Vlamakis H, *et al.* Compositional and Temporal Changes in the Gut Microbiome of Pediatric Ulcerative Colitis Patients Are Linked to Disease Course. *Cell Host Microbe* 2018;24:600–10.
- 54 Zhu Q, Hou Q, Huang S, *et al.* Compositional and genetic alterations in Graves' disease gut microbiome reveal specific diagnostic biomarkers. *ISME J* 2021;15:3399–411.
- 55 Wen C, Zheng Z, Shao T, *et al.* Quantitative metagenomics reveals unique gut microbiome biomarkers in ankylosing spondylitis. *Genome Biol* 2017;18:142.
- 56 Liu P, Jiang Y, Gu S, *et al.* Metagenome-wide association study of gut microbiome revealed potential microbial marker set for diagnosis of pediatric myasthenia gravis. *BMC Med* 2021;19:159.
- 57 Kong C, Liang L, Liu G, *et al.* Integrated metagenomic and metabolomic analysis reveals distinct gut-microbiome-derived phenotypes in early-onset colorectal cancer. *Gut* 2023;72:1129–42.
- 58 Song M, Chan AT, Sun J. Influence of the Gut Microbiome, Diet, and Environment on Risk of Colorectal Cancer. *Gastroenterology* 2020;158:322–40.
- 59 Shah SC, Itzkowitz SH. Colorectal Cancer in Inflammatory Bowel Disease: Mechanisms and Management. *Gastroenterology* 2022;162:715–30.
- 60 Thomas AM, Manghi P, Asnicar F, *et al.* Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *N Med* 2019;25:667–78.
- 61 Lee KA, Thomas AM, Bolte LA, *et al.* Cross-cohort gut microbiome associations with immune checkpoint inhibitor response in advanced melanoma. *N Med* 2022;28:535–44.
- 62 Zhao L, Zhang X, Zhou Y, *et al.* Parvimonas micra promotes colorectal tumorigenesis and is associated with prognosis of colorectal cancer patients. *Oncogene* 2022;41:4200–10.
- 63 Kostic AD, Chun E, Robertson L, *et al.* *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe* 2013;14:207–15.
- 64 Kabeerdoss J, Sankaran V, Pugazhendhi S, *et al.* *Clostridium leptum* group bacteria abundance and diversity in the fecal microbiota of patients with inflammatory bowel disease: a case-control study in India. *BMC Gastroenterol* 2013;13:20.
- 65 Grenda T, Grenda A, Domaradzki P, *et al.* Probiotic Potential of *Clostridium* spp.—Advantages and Doubts. *CIMB* 2022;44:3118–30.
- 66 Kim CH. Immune regulation by microbiome metabolites. *Immunology* 2018;154:220–9.
- 67 Zheng D, Liwinski T, Elinav E. Interaction between microbiota and immunity in health and disease. *Cell Res* 2020;30:492–506.
- 68 Wardman JF, Bains RK, Rahfeld P, *et al.* Carbohydrate-active enzymes (CAZymes) in the gut microbiome. *Nat Rev Microbiol* 2022;20:542–56.
- 69 Onyango SO, Juma J, De Paepe K, *et al.* Oral and Gut Microbial Carbohydrate-Active Enzymes Landscape in Health and Disease. *Front Microbiol* 2021;12:653448.
- 70 El Kaoutari A, Armougom F, Gordon JJ, *et al.* The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat Rev Microbiol* 2013;11:497–504.
- 71 Das UN. Pyruvate is an endogenous anti-inflammatory and anti-oxidant molecule. *Med Sci Monit Int Med J Exp Clin Res* 2006;12:RA79–84.
- 72 Yang R, Zhu S, Tonnessen TI. Ethyl pyruvate is a novel anti-inflammatory agent to treat multiple inflammatory organ injuries. *J Inflamm (Lond)* 2016;13:37.
- 73 Luu M, Riester Z, Baldrich A, *et al.* Microbial short-chain fatty acids modulate CD8⁺ T cell responses and improve adoptive immunotherapy for cancer. *Nat Commun* 2021;12:4077.
- 74 Furusawa Y, Obata Y, Fukuda S, *et al.* Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nat New Biol* 2013;504:446–50.
- 75 Nastasi C, Fredholm S, Willerslev-Olsen A, *et al.* Butyrate and propionate inhibit antigen-specific CD8⁺ T cell activation by suppressing IL-12 production by antigen-presenting cells. *Sci Rep* 2017;7:14516.
- 76 Deleu S, Machiels K, Raes J, *et al.* Short chain fatty acids and its producing organisms: An overlooked therapy for IBD? *EBioMedicine* 2021;66:103293.
- 77 Portincasa P, Bonfrate L, Vacca M, *et al.* Gut Microbiota and Short Chain Fatty Acids: Implications in Glucose Homeostasis. *Int J Mol Sci* 2022;23:1105.
- 78 Frampton J, Murphy KG, Frost G, *et al.* Short-chain fatty acids as potential regulators of skeletal muscle metabolism and function. *Nat Metab* 2020;2:840–8.
- 79 Wu Y, Wang CZ, Wan JY, *et al.* Dissecting the Interplay Mechanism between Epigenetics and Gut Microbiota: Health Maintenance and Disease Prevention. *IJMS* 2021;22:6933.
- 80 Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, *et al.* The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res* 2020;48:D845–55.
- 81 Krämer A, Green J, Pollard J, *et al.* Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* 2014;30:523–30.
- 82 Ueno H. The IL-12-STAT4 axis in the pathogenesis of human systemic lupus erythematosus. *Eur J Immunol* 2020;50:10–6.
- 83 Fang S-Y, Li C-L, Liu X-S, *et al.* Correlation between polymorphisms of the NR3C1 gene and glucocorticoid effectiveness in patients with pemphigus vulgaris. *Sci Rep* 2017;7:11890.
- 84 Xie F, Li G, Wang Y, *et al.* Pyridoxal phosphate synthases PdxS/PdxT are required for *Actinobacillus pleuropneumoniae* viability, stress tolerance and virulence. *PLoS ONE* 2017;12:e0176374.
- 85 Zhu Y, Yang S, Zhao N, *et al.* CXCL8 chemokine in ulcerative colitis. *Biomed Pharmacother* 2021;138:111427.

- 86 Mao Y-M, Zhao C-N, Liu L-N, *et al.* Increased circulating interleukin-8 levels in systemic lupus erythematosus patients: a meta-analysis. *Biomark Med* 2018;12:1291–302.
- 87 Ghafouri-Fard S, Shahir M, Taheri M, *et al.* A review on the role of chemokines in the pathogenesis of systemic lupus erythematosus. *Cytokine* 2021;146.
- 88 Chowers Y, Sela BA, Holland R, *et al.* Increased levels of homocysteine in patients with Crohn's disease are related to folate levels. *Am J Gastroenterol* 2000;95:3498–502.
- 89 Petta I, Dejager L, Ballegeer M, *et al.* The Interactome of the Glucocorticoid Receptor and Its Influence on the Actions of Glucocorticoids in Combatting Inflammatory and Infectious Diseases. *Microbiol Mol Biol Rev* 2016;80:495–522.
- 90 Katsanos KH, Voulgari PV, Tsianos EV. Inflammatory bowel disease and lupus: a systematic review of the literature. *J Crohns Colitis* 2012;6:735–42.
- 91 Yang Y, Nguyen M, Khetrapal V, *et al.* Within-host evolution of a gut pathobiont facilitates liver translocation. *Nat New Biol* 2022;607:563–70.
- 92 Bashford-Rogers RJM, Bergamaschi L, McKinney EF, *et al.* Analysis of the B cell receptor repertoire in six immune-mediated diseases. *Nat New Biol* 2019;574:122–6.
- 93 Schickel J-N, Glauzy S, Ng Y-S, *et al.* Self-reactive VH4-34-expressing IgG B cells recognize commensal bacteria. *J Exp Med* 2017;214:1991–2003.
- 94 Mavragani CP, Nezos A, Dovrolis N, *et al.* Type I and II Interferon Signatures Can Predict the Response to Anti-TNF Agents in Inflammatory Bowel Disease Patients: Involvement of the Microbiota. *Inflamm Bowel Dis* 2020;26:1543–53.
- 95 Postal M, Vivaldo JF, Fernandez-Ruiz R, *et al.* Type I interferon in the pathogenesis of systemic lupus erythematosus. *Curr Opin Immunol* 2020;67:87–94.
- 96 Schaupp L, Muth S, Rogell L, *et al.* Microbiota-Induced Type I Interferons Instruct a Poised Basal State of Dendritic Cells. *Cell* 2020;181:1080–96.
- 97 Gill SK, Rossi M, Bajka B, *et al.* Dietary fibre in gastrointestinal health and disease. *Nat Rev Gastroenterol Hepatol* 2021;18:101–16.
- 98 Wolter M, Grant ET, Boudaud M, *et al.* Leveraging diet to engineer the gut microbiome. *Nat Rev Gastroenterol Hepatol* 2021;18:885–902.
- 99 Kim CH. Complex regulatory effects of gut microbial short-chain fatty acids on immune tolerance and autoimmunity. *Cell Mol Immunol* 2023;20:341–50.
- 100 Zhang W, Mackay CR, Gershwin ME. Immunomodulatory Effects of Microbiota-Derived Short-Chain Fatty Acids in Autoimmune Liver Diseases. *J Immunol* 2023;210:1629–39.
- 101 Olson WJ, Jakic B, Hermann-Kleiter N. Regulation of the germinal center response by nuclear receptors and implications for autoimmune diseases. *FEBS J* 2020;287:2866–90.
- 102 McCurry MD, D'Agostino GD, Walsh JT, *et al.* Gut bacteria convert glucocorticoids into progestins in the presence of hydrogen gas. *Cell* 2024;187:2952–68.
- 103 Mukherji A, Kobiita A, Ye T, *et al.* Homeostasis in intestinal epithelium is orchestrated by the circadian clock and microbiota cues transduced by TLRs. *Cell* 2013;153:812–27.
- 104 Zhou H, Brito I. Data from: lupus and inflammatory bowel disease share a common set of microbiome features distinct from other autoimmune disorders. Zenodo; 2024. Available: <https://doi.org/10.5281/zenodo.13624416>