nature microbiology

ARTICLES

https://doi.org/10.1038/s41564-018-0306-4

Corrected: Author Correction

Gut microbiome structure and metabolic activity in inflammatory bowel disease

Eric A. Franzosa[1,2,11], Alexandra Sirota-Madi[1,11], Julian Avila-Pacheco[1], Nadine Fornelos[1], Henry J. Haiser[3], Stefan Reinker[3], Tommi Vatanen[1], A. Brantley Hall[1], Himel Mallick[1,2], Lauren J. McIver[1,2], Jenny S. Sauk[4], Robin G. Wilson[4], Betsy W. Stevens[4], Justin M. Scott[1], Kerry Pierce[1], Amy A. Deik[1], Kevin Bullock[1], Floris Imhann[5,6], Jeffrey A. Porter[7], Alexandra Zhernakova[6], Jingyuan Fu[6,8], Rinse K. Weersma[5], Cisca Wijmenga[6,9], Clary B. Clish[1], Hera Vlamakis[1], Curtis Huttenhower[1,2*] and Ramnik J. Xavier[1,4,10*]

The inflammatory bowel diseases (IBDs), which include Crohn's disease (CD) and ulcerative colitis (UC), are multifactorial chronic conditions of the gastrointestinal tract. While IBD has been associated with dramatic changes in the gut microbiota, changes in the gut metabolome—the molecular interface between host and microbiota—are less well understood. To address this gap, we performed untargeted metabolomic and shotgun metagenomic profiling of cross-sectional stool samples from discovery ($n = 155$) and validation ($n = 65$) cohorts of CD, UC and non-IBD control patients. Metabolomic and metagenomic profiles were broadly correlated with faecal calprotectin levels (a measure of gut inflammation). Across >8,000 measured metabolite features, we identified chemicals and chemical classes that were differentially abundant in IBD, including enrichments for sphingolipids and bile acids, and depletions for triacylglycerols and tetrapyrroles. While > 50% of differentially abundant metabolite features were uncharacterized, many could be assigned putative roles through metabolomic 'guilt by association' (covariation with known metabolites). Differentially abundant species and functions from the metagenomic profiles reflected adaptation to oxidative stress in the IBD gut, and were individually consistent with previous findings. Integrating these data, however, we identified 122 robust associations between differentially abundant species and well-characterized differentially abundant metabolites, indicating possible mechanistic relationships that are perturbed in IBD. Finally, we found that metabolome- and metagenome-based classifiers of IBD status were highly accurate and, like the vast majority of individual trends, generalized well to the independent validation cohort. Our findings thus provide an improved understanding of perturbations of the microbiome–metabolome interface in IBD, including identification of many potential diagnostic and therapeutic targets.

Inflammatory bowel disease (IBD) is a chronic inflammatory condition of the gastrointestinal tract that results from altered interactions between gut microbes and the intestinal immune system[1,2]. There are two main IBD subtypes, ulcerative colitis (UC) and Crohn's disease (CD), which localize in the large and small intestines, respectively, and are characterized by unique microbial signatures[3]. Previous studies have shown major shifts in the gut microbial composition of patients with IBD[2,4–8]. Likewise, microbial composition can shape the environment in the colon as metabolites that microbes produce can be involved in signalling, immune system modulation or have antibiotic activity[9–11]. However, it is less clear how specific microbes and the small molecules they modulate may interact to cause, sustain, mitigate or predict inflammatory conditions such as IBD.

Broadly, gut metabolite profiles are jointly derived from diet, modified human metabolites and microbially derived compounds that shape the microbiota–host interactions[9]. For example, short-chain fatty acids (SCFAs) such as butyrate, acetate and propionate are produced by gut bacteria when they break down dietary fibre. SCFAs can affect host cells by modulating histone deacetylase inhibitory activity, gene expression, cell proliferation and immune response[7,12]. In addition, butyrate can protect against colitis by regulating T_{reg} cell production and enhancing the antibacterial activity of macrophages[13,14]. In stool from patients with IBD there is a decrease in butyrate, an SCFA that is important in modulating the immune system along with a decrease in butyrate-producing bacteria[2,15].

Commensal microbes can also alter pools of available metabolites thereby modifying host-generated signalling molecules. Untargeted serum metabolomics of germ-free versus conventional mice showed that a large number of serum metabolites arise due to commensal microbes[16]. For example, tryptophan metabolism is largely affected by the presence of gut bacteria, since microbial tryptophan decarboxylases (among other enzymes) convert tryptophan from the diet into tryptamine and other molecules. Microbially derived

[1]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [2]Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA. [3]Novartis Institute for Biomedical Research Inc., Cambridge, MA, USA. [4]Gastrointestinal Unit and Center for the Study of Inflammatory Bowel Disease, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA. [5]Department of Gastroenterology and Hepatology, University of Groningen and University Medical Center Groningen, Groningen, The Netherlands. [6]Department of Genetics, University of Groningen and University Medical Center Groningen, Groningen, The Netherlands. [7]Novartis Institute for Biomedical Research Inc., Basel, Switzerland. [8]Department of Pediatrics, University of Groningen and University Medical Center Groningen, Groningen, The Netherlands. [9]Department of Immunology, K.G. Jebsen Coeliac Disease Research Centre, University of Oslo, Oslo, Norway. [10]Center for Microbiome Informatics and Therapeutics, Massachusetts Institute of Technology, Cambridge, MA, USA. [11]These authors contributed equally: Eric A. Franzosa, Alexandra Sirota-Madi. *e-mail: chuttenh@hsph.harvard.edu; xavier@molbio.mgh.harvard.edu

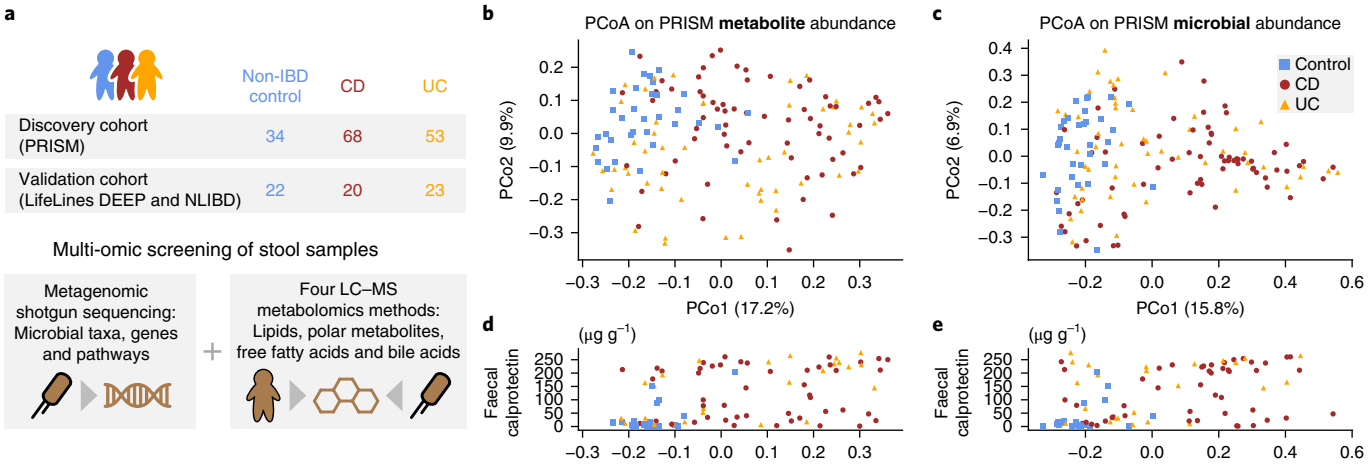

Fig. 1 | IBD is associated with broad changes in the gut multi-omic profiles of individuals. a, We collected and profiled stool metagenomic and metabolomic data from two IBD cohorts: a 155-member discovery cohort (PRISM) and a 65-member validation cohort (LifeLines DEEP and NLIBD). **b**, Principal coordinates analysis (PCoA) of PRISM cohort individuals based on gut metabolomic profiles (Bray–Curtis distance). **c**, The same individuals ordinated on Bray–Curtis distances between gut metagenomic species profiles. **d**,**e**, Patient faecal calprotectin levels (μg g^{-1}) plotted against the first PCoA axes from panels **b** and **c**, respectively. Note that faecal calprotectin measurements were not available for all individuals.

tryptophan metabolites alter host physiology not only by decreasing the available tryptophan, which can in turn perturb serotonin production, and by extension, behaviour[17], but also by producing indole derivatives that activate the aryl hydrocarbon receptor[18]. There is a decrease in tryptophan metabolism genes in microbiome samples from patients with CD[4]. A mouse study recently found that animals lacking one of the IBD susceptibility genes, *CARD9*, had altered microbial metabolism of tryptophan and were more susceptible to colitis[19].

A number of previous studies have identified differences in faecal metabolites in IBD[20–25]. However, these studies have tended to rely on small cohorts or 16S ribosomal RNA amplicon-based profiles of the associated IBD microbiota (that is, lacking shotgun metagenomic information). In a study of healthy individuals, untargeted faecal metabolomics correlated better with 16S-based microbiome composition than targeted metabolomics[26]. IBD-associated taxa were also highly correlated with metabotype in a study of inactive paediatric IBD. In the same study, healthy first-degree relatives displayed a similar microbiome and metabotype as relatives with inactive disease[25]. Both CD and UC gut microbiomes exhibit general decreases in taxonomic diversity relative to healthy gut microbiomes, along with phylum-level decreases in Firmicutes and increases in Proteobacteria[3,25,27]. In CD specifically, proportions of the Clostridia class are altered: the *Roseburia* and *Faecalibacterium* genera of the Lachnospiraceae and Ruminococcaceae families are decreased, whereas *Ruminococcus gnavus* increases[5,28,29]. Together, these findings suggest that yet-to-be characterized molecules in the gut metabolome, linked to inflammation and ultimately IBD, may be largely microbially derived or modified.

In this work, we took an unbiased approach to identify gut metabolites, microbial species and microbial enzymes that were differentially abundant in IBD relative to non-IBD controls. To that end, we performed untargeted liquid chromatography–mass spectrometry (LC–MS) metabolomic profiling and shotgun metagenomic sequencing of stool samples from a 155-member discovery cohort and a 65-member validation cohort, each containing a cross-sectional sampling of UC, CD and control patients. While metagenomic findings were largely in agreement with previous studies, metabolomic profiles revealed >2,700 differentially abundant metabolites in IBD, including 224 that were significantly elevated in both UC and CD. IBD-elevated metabolites were enriched for sphingolipids and bile acids (among other chemical classes), as well as many uncharacterized metabolites of potential microbial origin. Indeed, many differentially abundant metabolites participated in robust associations with differentially abundant microbial species and enzymes, which is suggestive of biological mechanisms relating their abundances. Finally, the vast majority of IBD associations from the discovery cohort were replicated in the independent validation cohort, thus making our findings a useful resource for the study of microbiome and metabolic perturbations in IBD.

Results

To characterize the gut metabolic profile and microbiome composition in IBD, we collected and analysed stool samples from a cross-sectional cohort of individuals enroled in PRISM (the Prospective Registry in IBD Study at MGH). This cohort included 155 patients: 68 with CD, 53 with UC and 34 non-IBD controls (Fig. 1a). Each stool sample was subjected to metagenomic sequencing followed by profiling of microbial community taxonomic composition and functional potential. In addition, each sample was analysed by four LC–MS methods measuring polar metabolites, lipids, free fatty acids and bile acids, respectively. LC–MS metabolomic profiling was carried out using sensitive, high-resolution mass spectrometers in non-targeted modes, thus capturing large numbers of known and uncharacterized metabolites, including those of potential microbial origin.

A total of 3,829 metabolite features (43% of total) were assigned to putative molecular classes based on comparisons with the Human Metabolome Database (HMDB)[30]; 466 features (representing 346 unique compounds) were annotated as standards through comparison with reference data generated from an in-house compound library (see Supplementary Datasets 1 and 2). Shotgun metagenomic and metabolomic data were then analysed (1) to identify IBD- and disease subtype-specific changes in individuals' microbial and metabolic profiles, (2) to describe associations between microbial and metabolite features and (3) to assess the power of these features to classify IBD status and subtype across populations. Relationships discovered in the PRISM cohort were validated against an independent cohort of 20 CD patients, 23 UC patients and 22 population controls from the Netherlands. These analyses are expanded in the following sections, with additional technical detail provided in the Methods.

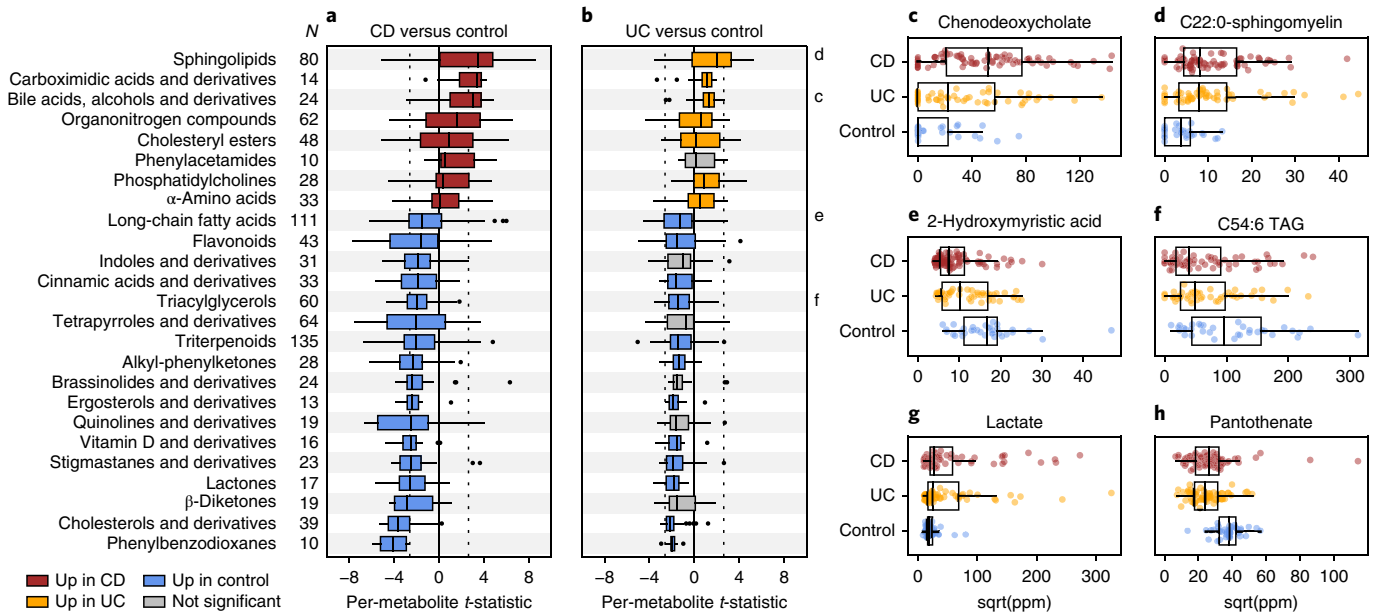

Fig. 2 | Metabolic enrichments in IBD versus control phenotypes. We applied Wilcoxon signed-rank tests to the individual differential abundance trends of metabolites (*t* statistics from the linear models) to identify classes of molecules that were broadly enriched in IBD. **a**, Focusing on classes of molecules with at least 10 putative members (see the *N* column), 8 were significantly (FDR $q < 0.05$) positively enriched in CD, meaning that their members tended to be more abundant in CD, and 17 classes were significantly negatively enriched, meaning that their members tended to be more abundant in controls (nominal *P* values were two-tailed). **b**, A subset of these trends was similarly significant in comparisons between UC and controls, with the remainder (grey) tending to trend in the same direction as CD versus control comparisons. The dotted line indicates the significance threshold for an individual metabolic feature (abs(*t*) > 2.61). **c**–**h**, Examples of individually differentially abundant standards measured across 68 CD, 53 UC and 34 non-IBD controls. Metabolites highlighted in panels **c**–**f** are representatives of broader classes analysed in **a** and **b**. Abundances are in ppm units after separately sum-normalizing within each LC–MS method; values are square-root-scaled for visualization. Boxplot 'boxes' indicate the first, second and third quartiles of the data. Boxplot 'whiskers' indicate the inner fences of the data, with points outside the inner fences plotted as outliers.

Broad metabolic shifts in IBD correlate with host inflammation. The major patterns of variation in the 155 PRISM patients' measurements of >8,000 metabolite features largely separated non-IBD controls versus CD patients, which is indicative of broad metabolic differences between these two phenotypes (Fig. 1b). Such differences could result from a combination of sources, including the effects of disease activity in host tissues, the activity of an IBD-altered microbiome and differences in patient diet and medication use. The metabolic profiles of UC patients were more broadly distributed, with roughly half resembling the profiles of non-IBD controls and the remainder more similar to the metabolomes of CD patients (Supplementary Fig. 1). Similar patterns of variation among disease phenotypes were apparent in the microbial taxonomic profiles from the patients' corresponding metagenomes (Fig. 1c; see also Supplementary Fig. 2 and Supplementary Dataset 4). Indeed, the first axes of ordination for the two datasets were well correlated (Spearman's $r = 0.664$, two-tailed $P < 10^{-20}$), consistent with strong coupling of gut metabolic profile, microbial community composition and disease status.

We hypothesized that broad variation in metabolic profile across individuals, especially within UC patients, might be explained in part by individuals' levels of active inflammation. We evaluated this by comparing the first axis of metabolic variation with individuals' levels of faecal calprotectin, a biomarker for severity of inflammation in IBD[31]. Across 93 individuals with faecal calprotectin measurements, the first axis of metabolic variation correlated in a reasonably strong and highly statistically significant manner with faecal calprotectin (Spearman's $r = 0.486$, two-tailed $P < 10^{-6}$; Fig. 1d). This correlation was driven in part by the tendency of control individuals to have very low faecal calprotectin levels (mean = 35 μg g^{-1}) and CD patients having very high levels (mean = 130 μg g^{-1}). However, the correlation remained strong and significant when evaluated on UC patients only ($n = 25$, $r = 0.565$, two-tailed $P = 0.003$). We observed a similar trend between faecal calprotectin measurements and the first axis of metagenomic variation (Fig. 1e), leading us to conclude that (1) our UC patients vary from control-like levels of inflammation to more active inflammation, and (2) that this variation may contribute to the more heterogeneous metabolic and metagenomic profiles of UC patients.

The first axes of metabolomic and taxonomic variation were also significantly associated with Shannon diversity (Supplementary Fig. 3). Consistent with previous findings, more inflamed, IBD-like samples (towards the right in Fig. 1c) tended to have markedly lower Shannon diversity (Spearman's $r = -0.572$, two-tailed $P < 10^{-14}$, $n = 155$). A similar, albeit weaker, trend was observed for metabolite profiles (Spearman's $r = -0.321$, $P < 10^{-4}$, $n = 155$), which exhibited less overall variation in within-sample diversity. Notably, these mutual associations with diversity were not sufficient to explain the strong coupling between the first axes of metabolomic and taxonomic variation, which remained significant after subtracting diversity effects using linear regression (residual correlation analysis, Spearman's $r = 0.364$, $P < 10^{-5}$, $n = 155$; see Methods).

The 68 CD patients in the PRISM (discovery) cohort were subclassified according to disease localization: L1 (ileal, $n = 14$); L2 (colonic, $n = 22$); L3 (ileocolonic, $n = 29$); L1 + L4 (ileal + upper gastrointestinal tract, $n = 1$); and unknown ($n = 2$). Compared to the strong separations we observed between CD and non-IBD individuals in the metabolomic and metagenomic data, we observed little to no stratification by disease localization among CD patients (Supplementary Fig. 4). More formally, overall diagnosis (CD/UC/

non-IBD) explained statistically significant fractions of the distance variation among individuals' metabolomic and metagenomic profiles (permutational analysis of variance, $P < 10^{-4}$; see Methods), while disease localization did not have a significant effect among CD patients ($P = 0.22$ and $P = 0.35$), possibly due to the established nature of IBD within the PRISM cohort. As a consequence of this finding, we treated CD as a single diagnosis in subsequent analyses.

Metabolite enrichments in IBD versus control phenotypes. To dissect metabolic changes in IBD at greater resolution, we applied a multivariable linear model to each metabolic feature to test association with IBD phenotype while controlling for other covariates (age and medication use; see Methods and Supplementary Dataset 3). Nominal *P* values for UC- and CD-specific effects were subjected to multiple hypothesis testing correction using the Benjamini-Hochberg[32] method with a false discovery rate (FDR) threshold of 0.05. Despite this strict filtering procedure, 2,729 metabolite features (31%) were significantly differentially abundant in IBD, including 200 matched against 151 unique standards. Out of all differentially abundant metabolites, the majority (1,931; 71%) were significantly depleted in IBD (CD or UC) relative to non-IBD controls; 224 (8%) were significantly elevated in both CD and UC; 505 (19%) were specifically elevated in CD; and only 69 (3%) were specifically elevated in UC (a possible consequence of the more heterogeneous metabolic profiles of UC patients). The large number of individually differentially abundant metabolites is consistent with the broad changes in metabolite profiles of IBD patients described earlier in the context of overview ordination (see Fig. 1b).

We performed enrichment analysis (Wilcoxon signed-rank tests; see Methods) to identify broad classes of compounds that were significantly over- or under-abundant in IBD phenotypes (ranking metabolite features by their CD- and UC-specific effect sizes). We defined metabolite classes based on HMDB annotations and focused on the 97 metabolite classes with at least 10 putative members in our dataset. Across these classes, we searched for enrichments in IBD or non-IBD controls that were statistically significant after correction for multiple hypothesis testing (Benjamini–Hochberg FDR $q < 0.05$). Eight of the 97 molecular classes were significantly overabundant in CD, with the strongest effects observed among sphingolipids, carboximidic acids and bile acids (Fig. 2a). Seven of these classes were additionally significantly overabundant in UC, while phenylacetamides were elevated, but not to a statistically significant degree (Fig. 2b). No molecular classes were specifically overabundant in UC.

IBD-enriched bile acids included cholate ($q = 0.003$) and chenodeoxycholate ($q = 0.0002$; Fig. 2c). In the healthy gut, these primary bile acids aid in the digestion of lipids and are deconjugated by microbes to secondary bile acids. We observed complementary depletions for the secondary bile acids sodium lithocholate and deoxycholate in CD, but the changes did not meet our threshold for FDR significance ($q = 0.06$ and 0.13, respectively). The relative overabundance of primary bile acids in the guts of IBD patients is consistent with disruption of bile acid transformation activities in the IBD microbiome[33]. Sphingolipids, another of the overabundant classes in IBD, play multiple roles in the healthy gut, including (1) as structural components of intestinal cell membranes and (2) as signalling molecules involved in cell fate decisions[34]. In addition to their presence in the membranes of human cells, sphingolipids are prevalent in the membranes of Bacteroidetes, and these microbially derived sphingolipids modulate the invariant natural T cell population[35]. Previous work suggested that sphingolipid metabolism may be disrupted in IBD, resulting in an accumulation of specific sphingolipid compounds that promote an inflammatory state[36–38]. Two of these compounds, ceramide and sphingomyelin (Fig. 2d), were significantly overabundant in both CD and UC ($q < 0.02$ in all comparisons).

Many more molecular classes were significantly depleted in CD and UC relative to controls (see Fig. 2a,b). Triterpenoids and long-chain fatty acids (LCFAs; including 2-hydroxymyristic acid, Fig. 2e) were the most numerous depleted classes (total $n = 135$ and 111, respectively), while phenylbenzodioxanes and cholesterols (including cholestenone) were the most consistently depleted classes. (The majority of their members were individually significantly depleted in CD.) Phenylbenzodioxanes are primarily derived from fruits, which reinforces the notion that some of the detected metabolic changes are a result of variation in individuals' diets. Triacylglycerols (TAGs), including C54:6 TAG (Fig. 2f), were additionally enriched in controls relative to CD and UC patients. This change, coupled with the enrichments for LCFAs and cholesterols, is consistent with previously suggested perturbations of fatty acid metabolism in IBD[39].

While their molecular classes were not generally differentially enriched in IBD, other notable differentially abundant metabolites included lactate (up in IBD; Fig. 2g) and pantothenate (down in IBD; Fig. 2h). Lactate has been previously reported as elevated in CD and UC patients[40], and is notable for being produced by members of the IBD gut microbiome, including lactobacilli, enterococci and pediococci. Pantothenate (vitamin B_5) is a precursor for coenzyme A, which is notable for being involved in fatty acid metabolism. Moreover, pantothenate is produced by the healthy gut microbiota; so (like lactate), its differential abundance in IBD patients may indicate a perturbation of microbe–metabolite relationships in the gut, a topic we explore in detail in a later section. While not statistically significant in this cohort, the SCFAs butyrate and propionate were decreased in both UC and CD patients relative to controls (Supplementary Dataset 3).

Modules of chemically related compounds are perturbed in IBD. To further explore biological patterns underlying the 2,729 differentially abundant metabolites, we clustered the differentially abundant metabolites based on the similarity of their residuals from the linear modelling approach described earlier. Metabolites co-clustered by this method will therefore tend to covary independently of their relationship with IBD phenotype, age and medication use. A total of 1,403 such clusters were identified with mean intracluster Spearman's $r > 0.7$. (Note that these unsupervised clusters, listed in Supplementary Dataset 1, are distinct from the HMBD-defined molecular classes used during enrichment analysis.) Each cluster was assigned a representative metabolite: the cluster centroid or the standard metabolite closest to the centroid (where applicable). The 50 largest clusters accounted for 780 differentially abundant features (29% of total), consistent with a smaller number of biological signals explaining many differentially abundant metabolites.

Clusters of covarying metabolites can arise by a variety of mechanisms, including: (1) chemical modification of a common parent metabolite; (2) metabolites interrelated by a biochemical pathway; (3) metabolites co-produced by a specific microbe; and (4) metabolites co-contributed from a specific dietary source. Biological signals suggested by metabolite covariation, especially those arising from interconversion of metabolites, can be used to transfer knowledge from annotated metabolites to their unannotated partners. This 'guilt-by-association' principle also arises in gene co-expression data, where it has been applied to identify modules of functionally related genes[41] and predict gene function assignments[42]. Following this logic, we found that co-clustered metabolites were 2.7× more similar in retention time, 3.0× more similar in mass/charge ratio and 15× more likely to belong to the same chemical class relative to random metabolite pairs (see Methods). Clusters are thus enriched for similar physico-chemical properties, and cluster co-membership may be predictive of such properties.

The largest metabolite cluster enriched in IBD (and the second largest overall) contained 39 metabolite features, all of them

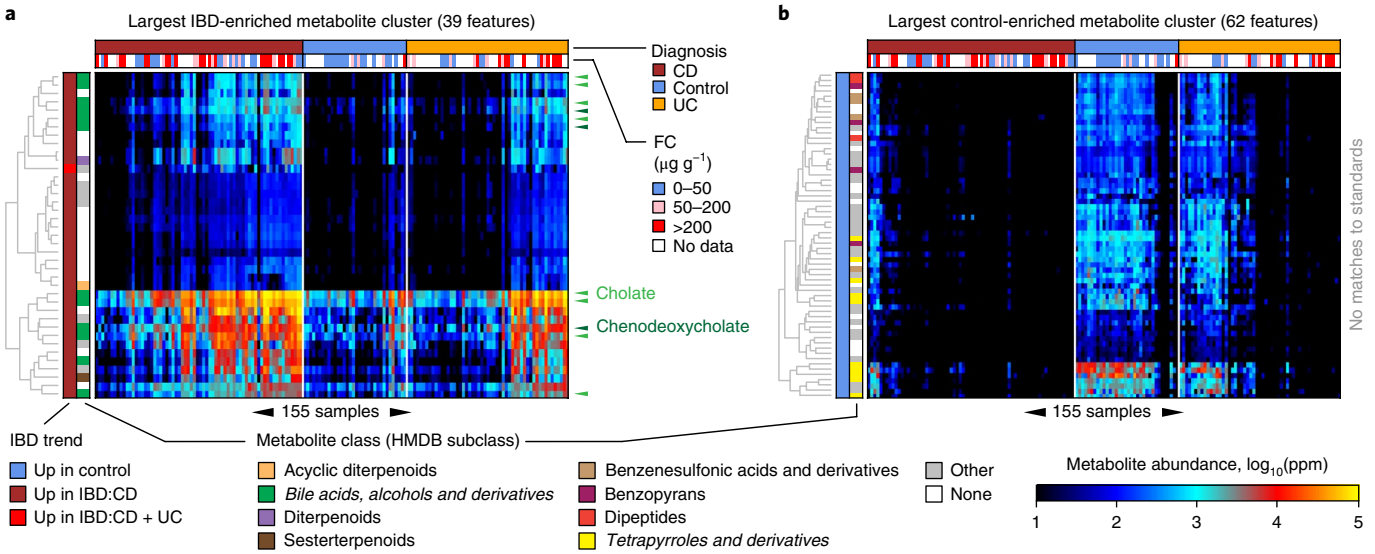

Fig. 3 | Clusters of chemically related, IBD-perturbed metabolites revealed by abundance covariation. We clustered differentially abundant metabolites after regressing out the effects of diagnosis, age and medication use (see Methods). A small number of (large) clusters explained many of the differentially abundant metabolites. **a**, The second largest cluster contained 39 metabolite features, all of them significantly elevated among CD patients (and one in UC patients also). This cluster was enriched for putative bile acids and derivatives. Multiple variants of the standards cholate (light green triangles) and chenodeoxycholate (dark green triangles) occur in this cluster. **b**, The largest cluster contained 62 metabolite features, all of them significantly elevated in non-IBD controls. This cluster was enriched for putative tetrapyrroles and derivatives. The 155 samples (columns) are ordered the same way in both panels according to Bray–Curtis similarity (and phenotype) of overall metabolic profile (as established in Supplementary Fig. 1). Note the control-like versus CD-like substructure among UC patients.

enriched in CD, with one additionally enriched in UC (Fig. 3a). This cluster contained 12 putative bile acids, including matches to cholate, chenodeoxycholate and their structural variants. This cluster also contained 17 unlabelled metabolites, which may also be related to bile acid metabolism via guilt-by-association logic. The largest cluster contained 62 metabolite features: all of them elevated among controls (Fig. 3b). Eleven features in this cluster were annotated as putative tetrapyrroles, but the cluster contained no validated standard metabolites, thus making it a promising target for further characterization. The previously described control-like and CD-like subdivisions of the UC population were also readily apparent in these individual clusters.

Other clusters of interest included the third largest cluster (33 members), which was consistently elevated among non-IBD controls and contained a variety of TAG metabolites (Supplementary Fig. 5). Cluster 13 (18 members) was uniquely elevated in CD patients and enriched for organonitrogen compounds, including the standards linoleoyl ethanolamide, palmitoylethanolamide and N-oleoylethanolamine (Supplementary Fig. 6). Clusters 23 and 25 were elevated in CD and UC patients and contained a variety of LCFAs, including the standards arachidonic acid, adrenic acid, docosapentaenoic acid (DPA) and eicosatrienoic acid (ETA) (Supplementary Fig. 7). Notably, the 99 clusters (7%) containing standards were more the exception than the rule; most clusters remain largely uncharacterized, allowing the potential for many previously undescribed, IBD-associated metabolites of microbial origin.

Species-level changes in IBD microbiome community composition. As introduced earlier, taxonomic profiling of individuals' gut microbiomes showed that the largest source of variation corresponded with the separation of non-IBD controls versus CD phenotypes, while UC patients were more heterogeneous (see Fig. 1b). To further dissect this trend, we applied the linear modelling approach introduced earlier to the abundances of 195 species-level clades (from 67 genera) that were present in at least five samples at 0.1% relative abundance (Supplementary Dataset 5). A total of 50 species were differentially abundant in one or more phenotypes, of which 35 were elevated in controls relative to IBD (Supplementary Fig. 2). *Roseburia hominis*, *Dorea formicigenerans* and *Ruminococcus obeum* were among the species exhibiting the strongest enrichments in non-IBD controls. The fact that these and many other species were significantly depleted in IBD relative to controls is consistent with the general trend towards loss of species diversity in the IBD microbiome[2,3,27] and with specific previous taxonomic enrichment studies[4,5,43,44]. Unclassified *Roseburia* species were significantly elevated in both CD and UC patients, while *Bifidobacterium breve* and *Clostridium symbiosum* were uniquely differentially abundant and enriched in UC. Twelve species were uniquely differentially abundant and enriched in CD, including *R. gnavus*, *Escherichia coli* and *Clostridium clostridioforme*. Many of these species-specific enrichments and depletions were in line with previous studies as cited and discussed in this article.

Putative mechanistic associations between IBD-linked microbes and metabolites. The multi-omic nature of this dataset enables the identification of microbial features and metabolites that (1) are mutually differentially abundant in IBD and (2) covary independently of their mutual covariation with disease. Such relationships are consistent with a mechanism relating the abundance of the species and metabolite that is then perturbed during IBD pathogenesis. For example, a positive association between a metabolite and species could indicate that the metabolite promotes the growth of that species, or that the species produces that metabolite. To identify such relationships, we performed large-scale association discovery between differentially abundant metabolites and species, focusing on representative differentially abundant metabolites and species from the clustering approach described earlier (notably, most species clustered alone by this approach). More importantly, we

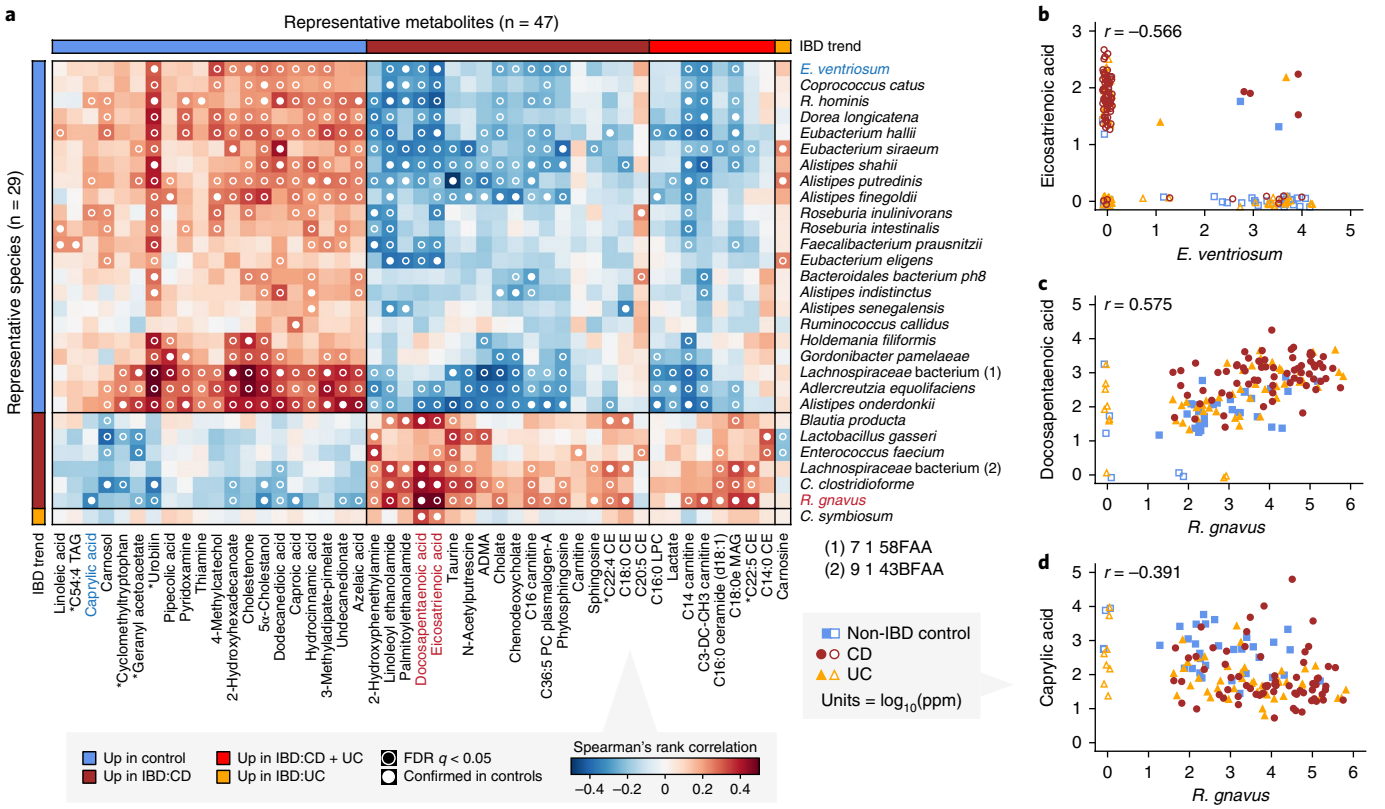

Fig. 4 | Potentially mechanistic associations between IBD-linked microbes and metabolites. a, Covariation between microbes and small molecules in IBD, specifically those linking FDR-significant, confirmed-in-controls metagenomic species and metabolites matched against standards (Spearman's rank correlation with two-tailed nominal *P* values). When multiple metabolomic features matched the same standard, the feature with the highest mean absolute correlation was selected for plotting. Metabolites marked with an asterisk indicate a match to a standard with isomeric forms that could not be differentiated. The standard L-1, 2, 3, 4-tetrahydro-beta-carboline-3-carboxylic acid is listed as 'cyclomethyltryptophan'. ADMA, asymmetric dimethylarginine; LPC, lysophosphatidylcholin; MAG, monoacylglycerol. **b**–**d**, Examples of individual correlations across 68 CD, 53 UC and 34 non-IBD controls (see text). The metabolites and species in these examples are coloured in panel **a**. Values plotted are raw measurements (not residuals) normalized to ppm units and then $\log_{10}$-transformed. Values < 1 ppm (including 0s) were set to 1 ppm for plotting; corresponding points are shown without fill and jittered; all other points have solid fill.

performed association discovery on metabolite and species residual abundances from the linear modelling approach described earlier, which will tend to de-emphasize associations driven purely by mutual association with disease status. This revealed a total of 15,679 FDR-significant ($q < 0.05$) associations between representative differentially abundant metabolites and species. Among these was a positive association between lactic acid and *Pediococcus acidilactici* (Spearman's $r = 0.23$), one of the expected microbe–metabolite relationships alluded to previously (see Fig. 2g). To further enrich for putatively mechanistic relationships that are perturbed in disease, we specifically focused on the subset of associations that were nominally significant ($P < 0.05$) and in the same direction when considering raw metabolite and species abundances from non-IBD controls only (we refer to these associations as 'confirmed in controls').

This filtered dataset encompassed 2,279 associations between differentially abundant metabolites and species (Supplementary Fig. 8), including 122 associations involving standards and characterized species (Fig. 4a). Associations covered 901 metabolite clusters representing 1,878 differentially abundant metabolites; 46 of 50 differentially abundant species were represented in at least one association. However, of the large number of possible associations between these metabolites and species, only 6% were statistically significant and confirmed in controls. This implies that, although many metabolites are associated with one or more species, they tend not to associate mechanistically with most species (and vice versa). The largest group of associations were positive associations between metabolites and species that were both elevated in controls (1,398 associations; 61% of all significant associations). These associations were representative of a general pattern of 'concordance' with disease, resulting, for example, when a species produces a protective metabolite. Discordant associations, for example, negative associations between metabolites and species that both increased in disease, accounted for only ~2% of total associations. In these cases, while the species and metabolite may be mechanistically linked, the mechanism does not appear to directly aggravate IBD pathogenesis.

The CD-associated compounds ETA and DPA were involved in negative associations with control-associated species and positive associations with IBD-associated species. ETA and DPA are polyunsaturated LCFAs and are examples of omega-3 and omega-6 fatty acids, respectively. ETA and DPA are important constituents of eukaryotic cell membranes; their elevation in the IBD-afflicted gut may be explained by higher rates of host cell death or turnover, or reduced absorption from diet. In addition to roles in immune and inflammatory signalling, polyunsaturated LCFAs possess bactericidal activity by virtue of their hydrophobic nature and potential to disrupt bacterial cell membranes[45]. This activity is particularly consistent with the negative correlations involving

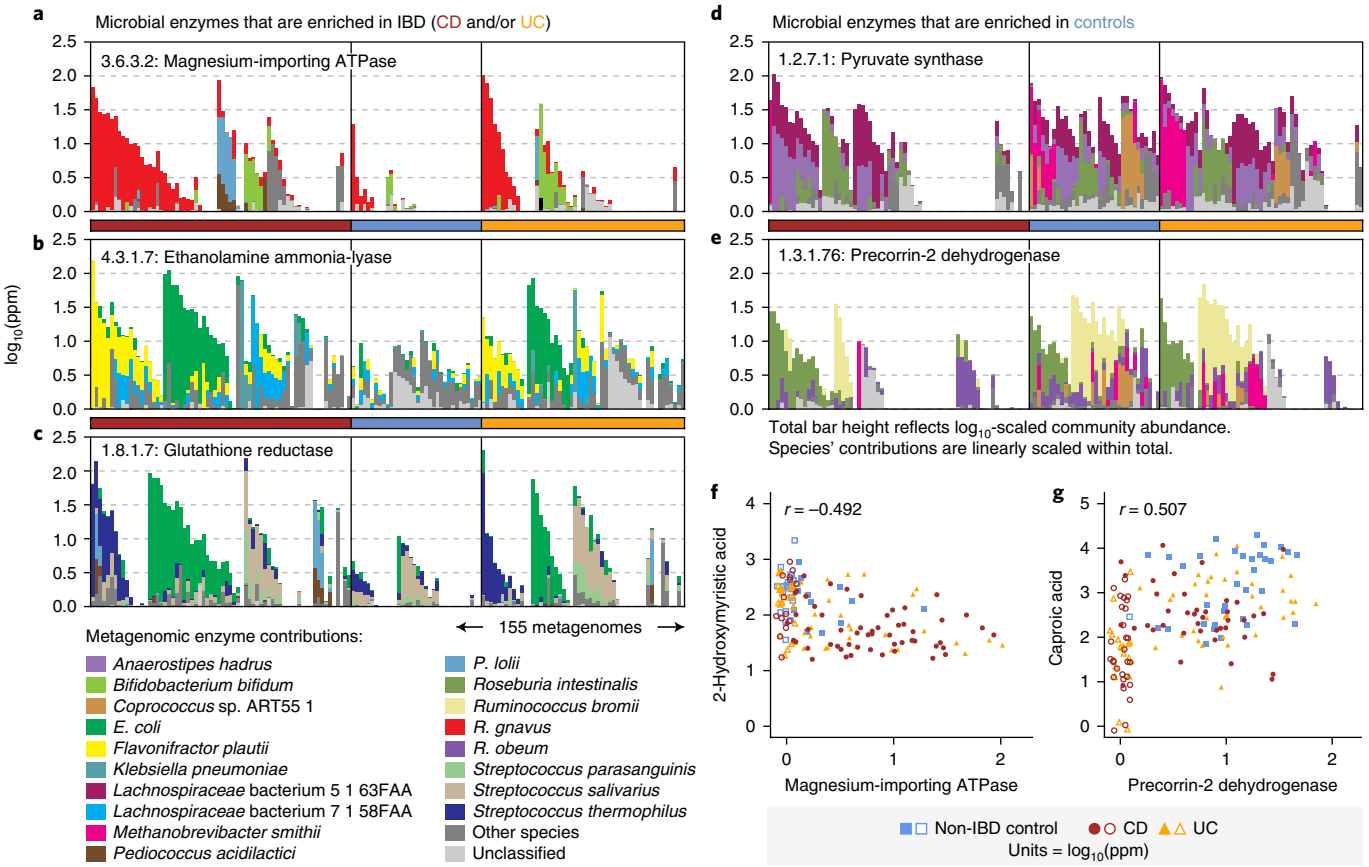

Fig. 5 | IBD-associated changes in microbial function and their metabolic associations. a–e, Examples of metagenomically contributed enzymes that were differentially abundant in IBD, annotated by their taxonomic contributors (**a–c** are enriched in IBD; **d** and **e** are depleted). In each case, the enzyme was contributed by a mixture of species across the cohort, and not dominated by a single species. Each set of stacked bars represents one of the 155 PRISM metagenomes (arrayed on horizontal axes). Community enzyme abundance (log_{10}-transformed ppm) is represented by the top of each stack of bars; contributions from major species are linearly scaled within the total bar height. Samples were first sorted according to the dominant contributor to a function and then grouped by phenotype (sample ordering differs between panels). **f**,**g**, Correlations between community total enzyme abundance and IBD-associated metabolites across 68 CD and 53 UC patients and 34 non-IBD controls. Values plotted are raw measurements (not residuals) normalized to ppm units and then log_{10}-transformed. Values < 1 ppm (including 0s) were set to 1 ppm for plotting; corresponding points are shown without fill and jittered (all other points have a solid fill). The given *r* values indicate Spearman's rank correlation.

ETA (for example, with *Eubacterium ventriosum*; Fig. 4b), several of which had a 'mutually exclusive' character (that is, when an ETA was present, the corresponding species tended to be absent, and vice versa). Conversely, DPA associated positively with IBD-associated species, most notably *R. gnavus* (Fig. 4c). This suggests that DPA encourages the growth of these species, possibly through disrupting the growth of health-associated species.

Caprylic acid, also known as octanoic acid, is a medium-chain fatty acid (MCFA) with antibacterial and antiviral properties[46]. Caprylic acid was enriched in non-IBD controls in our dataset, consistent with previously observed patterns of MCFA depletion in IBD[22]. Like SCFAs, MCFAs may occur in the gut as a breakdown product from anaerobic fermentation of fibre, although dietary contributions are perhaps more abundant. Consistent with this idea, caprylic acid was (weakly) positively correlated with a number of health-associated gut anaerobes, including *Alistipes shahii*, *Alistipes putredinis* and *Alistipes finegoldii*. On the other hand, caprylic acid was significantly negatively associated with the abundance of *R. gnavus* (Fig. 4d). Such a negative relationship would be consistent with possible uptake and metabolism of caprylic acid by *R. gnavus* (in which case, as the species abundance increases, more caprylic acid is used up). Alternatively, and more consistent with its aforementioned antibacterial properties, caprylic acid may have an inhibitory effect on the growth of *R. gnavus*.

To experimentally validate the potential for IBD-associated metabolites to exert growth effects on an IBD-associated species, we cultured *R. gnavus* in the presence of eight molecules with which it was observed to associate in the preceding analysis (see Methods). Among four predicted negative associations, caprylic acid indeed inhibited the growth of *R. gnavus* at high concentrations, as hypothesized above (Supplementary Fig. 9). Among four predicted positive associations, taurine and DPA were confirmed to enhance growth, while phytosphingosine exhibited a paradoxical inhibitory effect. Given the many factors that could impact the results of growth assays, for example, strain specificity and molecular concentrations, and the potential for mechanisms of association beyond direct effects on growth, for example, production as by-product, these results provide promising initial support for the usefulness of our multi-omic association framework in focusing downstream experiments.

IBD-associated changes in microbial function and their metabolic associations. To understand the functional consequences of microbial community changes in IBD, we first functionally

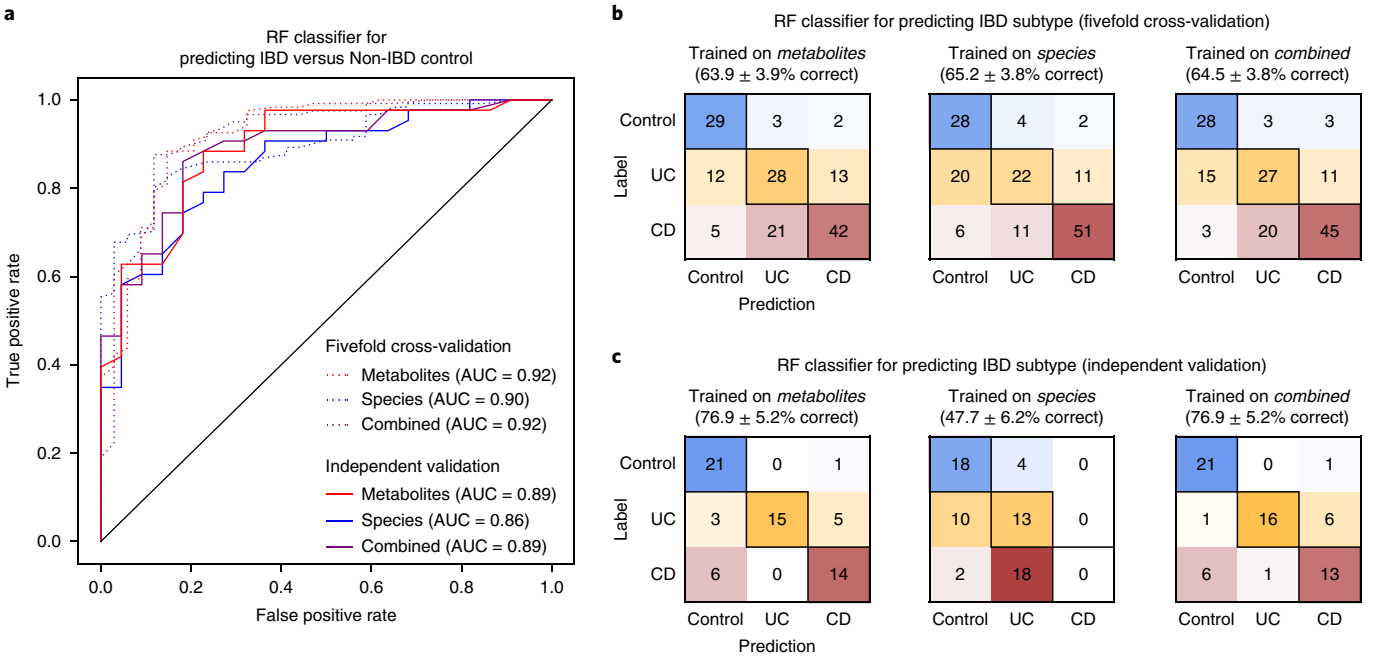

Fig. 6 | Predicting IBD status and subtype from gut microbiome multi-omic features. We trained RF classifiers on metabolites, microbial species and their combination to identify IBD patients and IBD subtypes. Training/testing was carried out within the PRISM cohort using fivefold cross-validation, in addition to models trained on the full PRISM cohort, and then tested (validated) on the independent Netherlands cohorts. **a**, Receiver operating characteristic (ROC) curves depict trade-offs between RF classifer true and false positive rates as classification stringency varies. The AUC statistic is a summary measure of classifier performance. AUC values close to 1 indicate that a high true positive rate was achieved with low false positive rate (ideal performance), while AUC values close to 0.5 indicate random performance. **b**, 'Confusion matrix' evaluations of IBD subtype RF classifiers within the PRISM cohort. The number in row *i* and column *j* indicates how many samples were labelled as subtype *i* but assigned to subtype *j*. A perfect subtype RF classifier (100% accuracy) would have 0 counts for all non-diagonal entries (that is, no misclassified samples). Matrix cells are shaded within-row in proportion to their value (red, CD; orange, UC; blue, non-IBD control). **c**, Confusion matrix evaluations of IBD subtype RF classifiers trained on the PRISM cohort and tested on the independent Netherlands cohorts. Accuracy values in **b** and **c** indicate the fraction of correctly classified instances; error values reflect the s.e.m of a proportion.

profiled gene families in all metagenomes using HUMAnN2, and then summed their abundances according to Enzyme Commission (EC) number annotations (see Methods and Supplementary Dataset 6). We applied the previously described linear modelling approach to this enzyme abundance data, revealing 568 enzymes that were differentially abundant (FDR-corrected $q < 0.05$) in CD, UC or both (Supplementary Fig. 10 and Supplementary Dataset 7). However, examining species-level functional attribution data, it was clear that many of these differentially abundant enzymes could be explained by a single IBD-associated species dominating contributions of the enzyme to the community. More specifically, when defining 'dominating' as 'explaining >50% of enzyme copies in >50% of samples', then *E. coli* alone dominated 220 differentially abundant enzymes, owing in part to that species' strong enrichment in IBD and exceptionally thorough functional annotations. While some enzymes in this category may indeed have mechanistic connections to IBD, others may simply expand (or shrink) in copy number alongside their source genomes, whose abundance is changing for reasons unrelated to encoding of that particular enzyme.

Of the differentially abundant enzymes, 246 were not dominated by any single species, suggesting that their enrichment in controls (or IBD) was better explained by a community-level shift in functional potential, and therefore of greater mechanistic significance. For example, magnesium-importing ATPase (EC 3.6.3.2) was enriched in both CD and UC patients relative to controls (Fig. 5a). Magnesium deficiency has been described as a known side effect of IBD[47], which could be explained in part by sequestration of the ion by the IBD-associated microbiome. Ethanolamine ammonia-lyase (EC 4.3.1.7) was similarly enriched in the IBD gut (Fig. 5b). This enzyme is involved in the production of glycerophospholipids, one of the most significantly enriched classes of metabolite in CD and UC patients (see Fig. 2a). A final example of an IBD-enriched enzyme was glutathione reductase (GR; EC 1.8.1.7; Fig. 5c). GR catalyses the production of glutathione, a compound involved in resistance to oxidative stresses. Oxidative stress is a hallmark of inflammation in the IBD-afflicted gut[48], thus giving species encoding GR a selective advantage in that environment.

Additional examples of differentially abundant enzymes were reflective of transitions from a more obligate to a facultative anaerobic microbiome in IBD. For example, pyruvate synthase (EC 1.2.7.1), an anaerobic enzyme that catalyses the interconversion of pyruvate and acetyl coenzyme A, was enriched in controls and completely undetected in a subpopulation of CD patients (Fig. 5d). Enzymes involved in the synthesis of cobalamin (vitamin B_{12}) were also enriched in controls, including precorrin-2 dehydrogenase (EC 1.3.1.76; Fig. 5e). While vitamin B_{12} (a tetrapyrrole-containing structure) is too large to be captured by the LC–MS methods used in this study, its derivatives and associated compounds may be among the putative tetrapyrroles that were enriched in the largest cluster of IBD-depleted metabolite features (see Fig. 3b).

To evaluate the potential mechanistic links between differentially abundant enzymes and metabolites more formally, we repeated the clustering and association procedures described earlier in the context of metabolite–species associations. Metabolite–enzyme associations followed many of the same patterns observed for species and enzymes. Association density was low (3%), suggesting that most

metabolites associated with only a few enzymes (and vice versa); the vast majority of interactions (95%) were concordant with IBD pathogenesis (Supplementary Fig. 11). Several associations occurred between standard metabolites and the enzymes discussed earlier. For example, magnesium-importing ATPase was strongly negatively associated with 2-hydroxymyristic acid (a control-enriched compound; Spearman's $r=-0.492$; Fig. 5f). Conversely, precorrin-2 dehydrogenase was positively associated with caproic acid (another control-enriched compound; $r=0.507$; Fig. 5g). While such relationships are consistent with compounds acting as enzyme substrates and products, respectively, this does not appear to be the case for these specific enzyme–compound pairs, suggesting that other factors probably mediate their associations (for example, encoding by or interaction with subsets of IBD-associated species).

Most IBD trends replicate in an independent validation cohort. We evaluated the generality of the differentially abundant metabolite features and microbial species identified earlier in an independent cohort of 20 CD and 23 UC patients and 22 non-IBD controls from the Netherlands (see Methods). Of 2,456 metabolite features that were differentially abundant in CD (up or down) in the PRISM cohort, 2,300 (94%) trended in the same direction in the Netherlands cohorts, of which 959 (39% of the grand total) were also FDR-significant (Supplementary Fig. 12). Of 1,049 metabolite features that were differentially abundant in UC (most of which were also differentially abundant in CD), 865 (82%) trended in the same direction, of which 117 (11% of the grand total) were FDR-significant. Similar patterns were observed for differentially abundant microbial species: 36 of 38 species that were differentially abundant in CD among PRISM patients trended in the same direction among Netherlands patients, with 13 achieving statistical significance. All 15 UC-significant species from the PRISM patients trended in the same direction among Netherlands patients, with 3 achieving statistical significance. Hence, the majority of IBD-associated changes identified in the PRISM cohort generalized in their directionality to the Netherlands cohorts. Statistical significance was not replicated as consistently, which we can attribute in part to loss of power in the Netherlands cohort from smaller sample size (total $n=65$ versus 155 for the PRISM cohort).

Multi-omic signatures differentiate IBD subtypes across cohorts. To evaluate if differences in metabolite or microbial composition could be used to classify patients according to IBD phenotype, we trained random forest (RF) classifiers on patient metabolic and microbial species profiles (separately and combined). Classification performance was evaluated within the PRISM cohort (using five-fold cross-validation) and between cohorts by training on the entire PRISM cohort and validating on the independent Netherlands cohort. In both of these approaches, RF classifiers are trained on one set of samples and then tested on another (non-overlapping) set, meaning that testing performance does not benefit from potential overfitting of RF classifiers to their training data.

All RF classifiers performed considerably better than random in the task of distinguishing IBD and non-IBD controls, with area under the curve (AUC) values ranging from 0.86 to 0.92 (AUC values close to 1.0 indicate that an classifier attained a high sensitivity at a very low false positive rate, while a value of 0.5 is expected at random; Fig. 6a and Supplementary Fig. 13). Cross-validation results (AUC 0.90–0.92) were only marginally better than independent validation results (AUC 0.86–0.89), indicating that the PRISM-trained RF classifier generalized well to the Netherlands cohort, which is consistent with the feature-level concordance described earlier. RF classifiers trained on metabolite features versus microbial species performed similarly, despite the metabolite feature space being considerably larger (thousands versus tens of features). The integration of metabolite and microbial species data did not produce a marked improvement in classification accuracy relative to metabolite features alone, which is consistent with a high degree of shared information between the gut's microbial and metabolomic profiles.

Predicting IBD subtype (summarized simply as CD and UC) was comparatively more challenging. Within the PRISM cohort, metabolites, species and their combination predicted UC, CD or non-IBD control labels correctly 64–65% of the time. This is less successful than case/control predictions, but still considerably greater than random (that is, 33% correct; Fig. 6b). The most common source of classification error was labelling UC patients as non-IBD or CD. This is not surprising, given that the distribution of UC patients overlapped with the (largely distinct) CD and non-IBD populations (see Fig. 1b). Comparatively, non-IBD individuals were rarely classified into one of the two IBD subtypes, while CD patients were sometimes erroneously classified as UC. More distinction among input data types was observed when applying the PRISM-trained IBD subtype RF classifier to the Netherlands cohorts (Fig. 6c). While the metabolite-incorporating RF classifiers performed reasonably well (77% correct classification), the species-based RF classifiers performed considerably worse (48% correct classification), largely because of marked misclassification of CD patients as UC. This suggests that, although many IBD-varying species trended similarly in the Netherlands cohort, subtype-informative details of their abundance distributions (as learned from the PRISM cohort) were less conserved.

Discussion

This study represents one of the first efforts to discover and validate IBD-associated changes in the human gut metabolome and microbiome in an integrated multi-omic framework. Many of the individually differentially abundant species and metabolic classes identified and validated in this study (for example, bile acids and sphingolipids) are in agreement with previous findings, while others (for example, dicarboxylic acids) are, to the best of our knowledge, unique to this study. More generally, we observed that metabolites and metabolite classes were frequently depleted in IBD patients relative to non-IBD controls. This pattern is suggestive of a loss of 'metabolic diversity' among IBD patients that is analogous to the loss of taxonomic (ecological) diversity observed in the IBD microbiome. This diversity is likely to be inclusive of a large number of previously undescribed, microbially derived metabolites that were unclassified or putatively classified in our comparisons with reference databases.

The ability of untargeted metabolomics approaches to quantify vast numbers of uncharacterized metabolites is both a strength and limitation relative to targeted approaches[26]. While uncharacterized metabolites no doubt encompass previously undescribed microbe- and disease-associated molecules of biological interest, they also include non-biological adducts and fragments of sample molecules, and are generally more challenging to interpret. We approached these challenges using a combination of methods: (1) experimentally validating metabolites against a standard compound library, a precise but resource-intensive process; (2) approximating annotation of metabolites to broad chemical classes; and (3) clustering of metabolites according to residual covariation across samples. Covariation-based clusters were found to be enriched for metabolites of similar physico-chemical properties, a form of guilt-by-association that complements existing network-based approaches[49] to metabolite characterization. The covariation-based approach suggested potential roles for many unlabelled metabolites that clustered with known standards (as in Fig. 3a). On the other hand, the process also revealed clusters of wholly uncharacterized, IBD-associated metabolites (as in Fig. 3b). Such clusters may represent microbial metabolites with pro-inflammatory or anti-inflammatory effects, and are prime candidates for additional experimental characterization.

Computational methods also provide a guide to downstream experimental validation and characterization of mechanisms relating the IBD microbiome and metabolome (including identification of microbiome-derived metabolites). By prioritizing associations between microbial species and enzymes that associate with metabolites independently of disease status, we enrich for potential mechanistic associations that may become perturbed in IBD. Many details of these associations remain to be determined. For example, a positive association between a microbial taxon and metabolite could be explained by (1) the metabolite representing a preferred carbon source that promotes species growth, (2) the metabolite occurring as a by-product of species metabolism, or (3) the metabolite selectively inhibiting the growth of other species (or otherwise interacting ecologically). These options can be disentangled computationally by analysing genome and metagenome annotations, when available, as well as experimentally by growing microbial species in the presence of their associated metabolites and/or profiling their metabolic output in monoculture. Naturally, further in vivo experiments (for example, in mouse models of IBD) are required to confirm that validated microbe–metabolite associations play a causal role in IBD pathogenesis. Such efforts are laborious; hence, the computationally derived subset of putative associations uncovered in this study will be a critical aid, as will further bioinformatic prioritization based on meta-omic profiling.

The vast majority of IBD-associated species and metabolites discovered in the PRISM cohort agreed in directionality with an independent validation cohort. Statistical significance was not as consistently replicated, partly because of the power limitations of the smaller validation study (hence, further replication in a larger cohort could be warranted in the future). At the same time, integrating individual microbial and metabolomic signals was sufficient to build accurate classifiers for case/control status that generalized to new individuals. Surprisingly, combining signals of both types (microbes and metabolites) did not boost classification performance markedly. This result is suggestive of tight coupling of the IBD gut metabolome and microbiome, which may result from a combination of (1) both profiles changing in response to disease, (2) an altered microbiome perturbing the metabolome or (3) an altered metabolome perturbing the microbiome (with potential feedback therein). Mechanisms underlying this coupling will be naturally expanded through experimental validation of targeted microbe–metabolite associations, as described earlier.

Predicting IBD subtype (UC versus CD) proved challenging, though this result was not surprising in light of other findings from the study. While CD patients did not stratify strongly by disease localization (see Supplementary Fig. 4), as a whole they separated well from non-IBD controls (see Fig. 1). The same could not be said for UC patients, which were dispersed into inflamed/CD-like and non-inflamed/control-like subpopulations, consistent with previous reports of high variability among UC microbiomes[25]. Many features that were individually differentially abundant in UC were also differentially abundant in CD, while the converse was not true. This is typical of the UC microbiome in general, and it suggests that IBD-linked perturbations may be divided into at least two modules: (1) perturbations that are associated with inflammation in general; and (2) perturbations that are specific to CD. The first module underlies the general association between multi-omic profiles and inflammation status (as measured in this study by faecal calprotectin level), and provides a basis for classifying case/control status. The second module can also aid in classifying case/control status, in that it is informative for CD subtype specifically. However, the absence of a strong UC-specific signal, coupled with heterogeneity among the UC subpopulation, hindered the predictability of UC status. That being said, a small number of molecules, including ethyl 9-hexadecenoate (see Supplementary Dataset 3), were individually differentially abundant in UC; they make promising targets for further study.

A number of future directions are possible for expanding this work to improve our understanding of metagenomic and metabolomic perturbations in IBD. To better differentiate UC, for example, it is possible that the metabolomic methods employed in this study missed classes of molecules that specifically vary with UC status. Alternatively, the stool metabolome may be imperfect for capturing UC-specific signals. In such cases, profiles of serum metabolites might augment serum antibodies[50] as diagnostic biomarkers for IBD/UC, while remaining less invasive than biopsy but still associable with the microbiome. While this study employed cross-sectional sampling of a larger number of individuals, dense longitudinal sampling of a subset of individuals would further aid in the dissection of putative microbe–metabolite associations (by intrinsically controlling for within-specific properties), and would further illuminate whether the observed population substructure was stable over time or correlated with changes in metabolomic or microbiome composition. Critically, this would also help disentangle causality—which metabolite shifts precede microbial or host phenotypes and vice versa—as well as provide a potential predictive target for interception of disease activity. However, even without these additional studies, the multi-omic screens and associations uncovered in this study provide many actionable hypotheses regarding the role of specific known and yet-to-be-characterized metabolites and their microbial partners in IBD pathogenesis. While many of these changes probably result from physiological changes on the host side, the subset that can be confirmed to result from microbial activity will provide promising targets for microbiome-based IBD diagnostics and therapies.

Methods

We performed untargeted metabolomic and metagenomic profiling on two IBD cohorts containing patients with CD and UC, and non-IBD controls. One cohort consisted of patients seen at the Massachusetts General Hospital (MGH, Boston, USA), and formed the basis of most analyses. A second (more heterogeneous) group of patients from the Netherlands was used to validate findings. Microbial species, microbial enzymes and >8,000 metabolites were tested for differential abundance in IBD. Differentially abundant metabolites were clustered to identify groups of functionally related compounds that were similarly perturbed in IBD. Differentially abundant metabolites and microbial features were compared using multi-omic correlation to identify putative mechanistic associations. Finally, all features were applied to build and validate multi-omic classifiers for IBD status and subtype.

PRISM cohort description and sample handling. PRISM is a referral centre-based, prospective cohort of IBD patients; 161 adult patients (>18 years old) enroled in PRISM and diagnosed with CD, UC and non-IBD (control) were selected for this study, with diagnoses based on standard endoscopic, radiographical and histological criteria. The PRISM research protocols were reviewed and approved by the Partners Human Research Committee (ref. 2004-P-001067), and all experiments adhered to the regulations of this review board. PRISM patient stool samples were collected at the MGH gastroenterology clinic and stored at −80 °C before DNA was extracted.

Validation cohort description and sample handling. The validation cohort consisted of 65 patients enroled in two distinct studies from the Netherlands; 22 controls were enroled in the LifeLines DEEP general population study[51] and 43 patients with IBD were enroled in a study at the Department of Gastroenterology and Hepatology at University Medical Center Groningen. Patients enroled in both studies collected stool using the same protocol: a single stool sample was collected at home and then frozen within 15 min in a conventional freezer. A research nurse visited all participants at home to collect home-frozen stool samples, which were then transported and stored at −80 °C. The stool samples were kept frozen before DNA was extracted or metabolomic profiling took place, as described in the ensuing sections.

Approval for human patient research. Human patient research in the discovery (PRISM) cohort was reviewed and approved by the Partners Human Research Committee (ref. 2004-P-001067), and all experiments adhered to the regulations of this review board. Human patient research in the validation cohorts (LifeLines DEEP and NLIBD) was approved by the University Medical Center Groningen review board (ref. M12.113965 and Institutional Review Board no. 2008.338, respectively). All study procedures were performed in compliance with all relevant ethical regulations for the validation cohorts. Each participant signed an informed consent form prior to participation for PRISM and both validation cohorts.

DNA extraction and metagenomic sequencing. Metagenomic data generation and processing were performed at the Broad Institute (Cambridge, USA). Stool DNA extractions were carried out using the QIAamp DNA Stool Mini Kit (QIAGEN). Whole-genome shotgun libraries were prepared by quantifying metagenomic DNA samples with the Quant-iT PicoGreen dsDNA Assay (Thermo Fisher Scientific) and normalized to a concentration of 50 pg µl^{-1}. Illumina sequencing libraries were prepared from 100–250 pg of DNA using the Nextera XT DNA Library Preparation Kit (Illumina) according to the manufacturer's recommended protocol, with reaction volumes scaled accordingly. Batches of 24, 48 or 96 libraries were pooled by transferring equal volumes of each library using an Echo 550 Liquid Handler (Labcyte). Insert sizes and concentrations for each pooled library were determined using an Agilent Bioanalyzer DNA 1000 kit (Agilent Technologies). Metagenomic libraries were sequenced on the HiSeq 2500 platform (Illumina), targeting ~2.5 Gb of sequence per sample with 101 base pair, paired-end reads.

Read-level quality control and metagenomic profiling. Raw sequencing reads were quality-controlled with KneadData version 0.5.1 (http://huttenhower.sph.harvard.edu/kneaddata). Briefly, this involved trimming low-quality bases from the 3′ end of reads with Trimmomatic[52] and then discarding trimmed reads <60 nt in length. Host (human) reads were identified and removed by mapping against the human genome (hg19 build) with Bowtie 2 (ref. [53]).

Quality-filtered metagenomes were taxonomically profiled using MetaPhlAn2 version 2.2.0 (ref. [54]) with default parameters. Only species-level relative abundance data were considered in this study. Species that failed to exceed 0.1% relative abundance in at least five samples were excluded. Functional profiling was performed using HUMAnN2 version 0.9.4 in UniRef90 mode (http://huttenhower.sph.harvard.edu/humann2)[55]. HUMAnN2 initially maps metagenomic reads to the pangenomes of species identified during taxonomic profiling (using Bowtie 2). Coding sequences in these pangenomes have been pre-annotated to their respective UniRef90 families[56]. Reads that did not align to a pangenome were mapped to UniRef90 by translated search with DIAMOND[57]. Hits to UniRef90-annotated sequences are weighted according to alignment quality, sequence length and sequence coverage. Gene-level outputs are produced in reads per kilobase units and stratified according to known/unclassified community contributions. Per-sample gene abundances were sum-normalized to ppm units. Gene abundances can be regrouped to other functional annotation systems based on annotations from UniProt[58]. For this study, gene abundances were regrouped (summed) according to EC number.

Metabolite profiling from stool samples. The gut metabolomic profiles of participants were measured from stool samples using a combination of four LC–MS methods that measure complementary metabolite classes. These range from polar metabolites (for example, organic acids), lipids (for example, triglycerides), free fatty acids and bile acids. In each method, the MS data were acquired using sensitive, high-resolution mass spectrometers (full scan MS profiling using Q Exactive Hydro Quadrupole Orbitrap and Exactive Plus Mass Spectrometers, Thermo Fisher Scientific) that enabled non-targeted measurement of (1) metabolites of known identity and (2) heretofore unidentified metabolites (for example, microbe-derived) in the same run.

Stool samples (weight range: 50.5–167.8 mg) were homogenized in 4 µl of water per mg stool sample weight using a bead mill (TissueLyser II; QIAGEN) and the aqueous homogenates were aliquoted for metabolite profiling analyses. Four separate LC–MS methods were used to measure polar metabolites and lipids in each sample. Methods 1, 2 and 3 were conducted using two LC–MS systems comprised of Nexera X2 U-HPLC System (Shimadzu Scientific Instruments) and Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer. Method 4 was conducted using a Nexera X2 U-HPLC coupled to an Exactive Plus Orbitrap Mass Spectrometer (Thermo Fisher Scientific).

Method 1: positive ion mode MS analyses of polar metabolites. LC–MS samples were prepared from stool homogenates (10 µl) via protein precipitation with the addition of nine volumes of 74.9:24.9:0.2 v/v/v acetonitrile/methanol/formic acid containing stable isotope-labelled internal standards (L-valine-d8, Isotec; L-phenylalanine-d8, Cambridge Isotope Laboratories). The samples were centrifuged (10 min, 9,000*g*, 4 °C), and the supernatants were injected directly onto a 150 × 2 mm Atlantis Silica HILIC Column (Waters). The column was eluted isocratically at a flow rate of 250 µl min^{-1} with 5% mobile phase A (10 mM ammonium formate and 0.1% formic acid in water) for 1 min followed by a linear gradient to 40% mobile phase B (acetonitrile with 0.1% formic acid) over 10 min. MS analyses were carried out using electrospray ionization in the positive ion mode using full scan analysis over *m/z* 70–800 at 70,000 resolution and 3 Hz data acquisition rate. Additional MS settings were: ion spray voltage, 3.5 kV; capillary temperature, 350 °C; probe heater temperature, 300 °C; sheath gas, 40; auxiliary gas, 15; S-lens RF level 40.

Method 2: negative ion mode MS analyses of polar metabolites. LC–MS samples were prepared from stool homogenates (30 µl) via protein precipitation with the addition of four volumes of 80% methanol containing inosine 15N4, thymine-D4 and glycocholate-D4 internal standards (Cambridge Isotope Laboratories). The samples were centrifuged (10 min, 9,000*g*, 4 °C) and the supernatants were injected directly onto a 150 × 2.0 mm Luna NH2 column (Phenomenex). The column was eluted at a flow rate of 400 µl min^{-1} with initial conditions of 10% mobile phase A (20 mM ammonium acetate and 20 mM ammonium hydroxide in water) and 90% mobile phase B (10 mM ammonium hydroxide in 75:25 v/v acetonitrile/methanol) followed by a 10 min linear gradient to 100% mobile phase A. MS analyses were carried out using electrospray ionization in the negative ion mode using full scan analysis over *m/z* 60–750 at 70,000 resolution and 3 Hz data acquisition rate. Additional MS settings were: ion spray voltage, 3.0 kV; capillary temperature, 350 °C; probe heater temperature, 325 °C; sheath gas, 55; auxiliary gas, 10; and S-lens RF level 40.

Method 3: negative ion mode analyses of metabolites of intermediate polarity (for example, bile acids and free fatty acids). Stool homogenates (30 µl) were extracted using 90 µl methanol containing prostaglandin E2-d4 as the internal standard (Cayman Chemical) and centrifuged (10 min, 9,000*g*, 4 °C). The supernatants (10 µl) were injected onto a 150 × 2 mm ACQUITY UPLC HSS T3 column (Waters). The column was eluted isocratically at a flow rate of 400 µl min^{-1} with 25% mobile phase A (0.1% formic acid in water) for 1 min followed by a linear gradient to 100% mobile phase B (acetonitrile with 0.1% formic acid) over 11 min. MS analyses were carried out using electrospray ionization in the negative ion mode using full scan analysis over *m/z* 200–550 at 70,000 resolution and 3 Hz data acquisition rate. Additional MS settings were: ion spray voltage, 3.5 kV; capillary temperature, 320 °C; probe heater temperature, 300 °C; sheath gas, 45; auxiliary gas, 10; and S-lens RF level 60.

Method 4: polar and non-polar lipids. Lipids were extracted from stool homogenates (10 µl) using 190 µl isopropanol containing 1-dodecanoyl-2-tridecanoyl-sn-glycero-3phosphocholine as the internal standard (Avanti Polar Lipids). After centrifugation (10 min, 9,000*g*, ambient temperature), supernatants (10 µl) were injected directly onto a 100 × 2.1 mm ACQUITY UPLC BEH C8 column (1.7 µm; Waters). The column was eluted at a flow rate of 450 µl min^{-1} isocratically for 1 min at 80% mobile phase A (95:5:0.1 v/v/v 10 mM ammonium acetate/methanol/acetic acid), followed by a linear gradient to 80% mobile phase B (99.9:0.1 v/v methanol/acetic acid) over 2 min, a linear gradient to 100% mobile phase B over 7 min and then 3 min at 100% mobile phase B. MS analyses were carried out using electrospray ionization in the positive ion mode using full scan analysis over *m/z* 200–1,100 at 70,000 resolution and 3 Hz data acquisition rate. Additional MS settings were: ion spray voltage, 3.0 kV; capillary temperature, 300 °C; probe heater temperature, 300 °C; sheath gas, 50; auxiliary gas, 15; and S-lens RF level 60.

Post-processing. We used Expressionist version 9.0 (Genedata; Refiner module for MS) to process raw LC–MS data for chemical noise removal, to detect chromatographic peaks and isotope clusters, align retention times between samples and assign putative metabolite identities via database look up. Detailed parameter settings are provided as Supplementary Dataset 8. Across samples, the combination of the four LC–MS methods generated 8,869 clustered features, characterized by chromatographic retention time and exact mass to < 5 ppm accuracy. Note that these clustered features, referred to as metabolites or metabolite features elsewhere in the text, are presumed to represent a single molecular species. Broader clusters of metabolite features, presumed to represent families of related molecular species, were also constructed using the results of linear regression analysis and are described later. Within each sample and LC–MS method, feature intensities were sum-normalized to ppm units.

A subset of 466 metabolites were identified more precisely using reference data generated from an in-house compound library; 3,829 metabolite features were linked to putative identifiers based on accurate *m/z* matching against the HMDB. Analyses of putatively matched features in the text focus on their molecular classes, rather than their identities. More specifically, we assigned HMDB subclasses to these features as a form of broad chemical classification. Subclasses assigned to >100 features (for example, 'fatty acyls') were further broken down according HMDB's 'direct parent' annotations.

Profile-level quality control. Before downstream analysis, metagenomic and metabolomic samples were subjected to profile-level quality control. First, we isolated the set of individuals with complete profiles of both types. All 65 Netherlands patients passed this filter, while 6 of 161 PRISM patients were missing one of the two profiles, for example, because of a failed sequencing run, and were excluded from subsequent analysis. Next, for both profile types, we considered the median Bray–Curtis distance of each PRISM sample to other samples in the PRISM cohort (the same distances form the basis of the ordinations in Fig. 1). If this distance was unusually large (defined as 'above the upper inner fence of all values'), the sample was considered an outlier. All PRISM metagenomic and metabolomic profiles passed this filter. Repeating this procedure within-phenotype (CD, UC, non-IBD control), we identified one potential control outlier among the PRISM metagenomic profiles, and a different control outlier plus one UC outlier among the PRISM metabolomic profiles. Because these profiles were representative of the human gut microbiome as a whole (if not their specific phenotype), they were retained for further analyses.

Statistical analyses. We carried out ordination analyses (Fig. 1; Supplementary Figs. 3 and 4) using classical multidimensional scaling on matrices of between-

sample diversity scores (Bray–Curtis distance). We used the Shannon diversity index to quantify within-sample diversity. Metabolomic diversity scores considered all measured metabolites (sum-normalized first within-method and then within-sample), while taxonomic diversity scores focused on species-level relative abundances. To control for within-sample diversity when comparing ordination axes, we generated best-fit lines between axis values and dataset-specific diversity measures, saved the resulting residual values and then compared dataset-specific residuals with Spearman's rank correlation. Other comparisons involving between-sample diversity and sample metadata were made using permutational analysis of variance as implemented in the 'adonis' function from R's 'vegan' package (using 10^4 permutations). Specifically, we computed the influence of diagnosis (CD, UC, non-IBD) across all patient metabolomic and metagenomic distances, and the influence of disease localization across the metabolomic and metagenomic profiles of CD patients. These analyses did not consider additional covariates.

We used linear models implemented in Python's 'statsmodels' package to identify microbial species, enzymes and metabolite features that were differentially abundant in IBD (http://www.statsmodels.org). Each data type was analysed separately in each cohort. Relative abundance values were log-transformed to variance-stabilize the data. Zero values were additively smoothed by half the smallest non-zero measurement on a per-sample basis. For both cohorts, we modelled the transformed abundance of each feature as a function of IBD phenotype (modelled as a categorical variable with 'non-IBD control' as the reference state), with age as a continuous covariate in both cohorts, and four medications (antibiotics, immunosuppressants, mesalamine and steroids) as binary covariates in the PRISM cohort. Effect sizes take the form of model *t* statistics (CD versus non-IBD control and UC versus non-IBD control) with associated two-tailed *P* values. Nominal *P* values were adjusted for multiple hypothesis testing with a target FDR of 0.05. A feature (metabolite, species or enzyme) was considered 'differentially abundant' in IBD if it passed this filter in either the CD- or UC-centred comparisons. Residual abundance values from the linear models were retained for use in subsequent analyses.

We identified molecular classes (as defined earlier) that were significantly enriched or depleted in IBD using rank-based enrichment analysis. Specifically, each metabolite was ranked according to its *t* statistics for CD- or UC-focused comparisons. For each class of molecule, we then evaluated if its members were enriched at the top or bottom of the list by performing a Wilcoxon signed-rank comparison of *t* values in the class versus those outside the class. Only classes with at least ten putative members were evaluated. Enrichment *P* values were corrected for multiple hypothesis testing as described earlier.

Unsupervised clustering. We performed clustering of differentially abundant features using a custom approach. Features were clustered on their residual abundance values from the linear modelling approach described earlier. This procedure enriches for covariation between features that is independent of mutual covariation with disease status (or other patient metadata, such as age or medication use). Features were ranked according to the significance of their association with IBD (the smaller of the two *P* values from the CD- and UC-centred comparisons). The highest-ranked feature was seeded into an initial cluster. Each subsequent feature was then compared to each extant cluster. If the feature had a mean similarity to the cluster's members exceeding a threshold, the feature was added to that cluster. (For all clustering analyses, we applied Spearman's rank correlation as a similarity measure with a threshold of $r=0.7$.) If the feature was not added to a cluster in this way, it was used to seed a new cluster. After considering all features, clusters were renumbered according to their size, such that cluster 1 had the most members, and so on. Each cluster was characterized by a representative member. For metabolite clusters containing standards, this representative was the standard closest to the cluster centroid; the true centroid was used for clusters without standards. Similarly, characterized (versus 'unclassified') species-level taxonomic features were preferred as representative features in microbial species clusters.

To evaluate guilt-by-association principles across the metabolite clusters, we compared pairs of metabolites present in the same cluster to all pairs of metabolites present in clusters with two or more members (that is, ignoring singleton clusters). To compare retention times, we evaluated the median difference in retention time for co-clustered versus all metabolite pairs (1.4 versus 3.8 min, a 2.7-fold reduction). The same procedure was used to compare mass/charge ratios (59 versus 174 Da, a threefold reduction). To compare chemical class, we restricted the analysis to annotated metabolites present in clusters with at least two annotated members; 18.8% of co-clustered metabolites were annotated to the same class, compared with 1.2% of all metabolites, a 15-fold enrichment for similarity.

Random forest classification. We performed random forest classification using the implementation of this method in Python's scikit-learn package (http://scikit-learn.org/). We considered separate RF classifiers for predicting (1) IBD/control status and (2) CD/UC/control status. We trained RF classifiers on the PRISM cohort using (1) fivefold cross-validation and (2) treating the entire cohort as a training set for independent validation against the Netherlands cohort. In each case, subject labels were randomly balanced before training and 100 trees were considered (other scikit-learn defaults were left unchanged). Features were not filtered in any way before random forest training (that is, the RF classifer could sample from any of the measured metabolites and/or species). Feature importance scores were retained for downstream analysis.

Growth effects of metabolites on *R. gnavus*. We grew *R. gnavus* (ATCC) in brain heart infusion medium medium (37 g l^{-1}) containing: 5% sterile-filtered foetal bovine serum (Sigma-Aldrich); 1% vitamin K_1-hemin solution (BD Biosciences); 1% trace mineral supplement (ATCC); 1% vitamin supplement (ATCC); 1 g l^{-1} D-(+)-cellobiose (Sigma-Aldrich); 1 g l^{-1} D-(+)-maltose (Sigma-Aldrich); 1 g l^{-1} D-(+)-fructose (Sigma-Aldrich); and 0.5 g l^{-1} L-cysteine (Sigma-Aldrich). Growth occurred under anaerobic conditions (atmosphere 5% H_2, 20% CO_2, 75% N_2) in a soft-sided vinyl chamber (Coy Laboratory Products). We sterilized the media using a Corning filter unit (0.22 µm pore diameter). All metabolite standards (Sigma-Aldrich) were brought to 100 mM in DMSO (Sigma-Aldrich) before dilution for dose assays. Overnight bacterial cultures were diluted 100-fold in appropriate media and 40 µl were dispensed per well in 384-well microplates (low evaporation lid, Costar 3680) containing metabolites or DMSO control. The microplates were shaken to ensure homogeneity; bacterial growth was monitored anaerobically (absorbance at 600 nm) in a microplate reader (PowerWave HT Microplate Spectrophotometer, BioTek) for 24 h at 37 °C without shaking. The values recorded for DMSO controls and metabolite-treated triplicates were averaged.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability. With the exception of Genedata Expressionist, the software packages used in this study are free and open source. The bioBakery tools (including KneadData, MetaPhlAn2 and HUMAnN2) are available via http://huttenhower.sph.harvard.edu/biobakery as source code and installable packages. The Python packages SciPy, Matplotlib (used for all data visualizations), statsmodels and scikit-learn are available from http://pypi.python.org. The R package vegan is available from http://cran.r-project.org. Analysis scripts employing these packages (and associated usage notes) are available from the authors upon request.

Data availability

Metagenomic sequences for the PRISM, LifeLines DEEP and NLIBD cohorts are available via SRA with BioProject number PRJNA400072. Metabolomics data (accession number PR000677) are available at the National Institutes of Health Common Fund's Metabolomics Data Repository and Coordinating Center (supported by National Institutes of Health grant no. U01-DK097430): Metabolomics Workbench (http://www.metabolomicsworkbench.org). Tables of processed metabolite, microbial species and microbial enzyme abundance are available as Supplementary Datasets 2, 4 and 6.

Received: 16 November 2017; Accepted: 25 October 2018;
Published online: 10 December 2018

References

1. Wlodarska, M., Kostic, A. D. & Xavier, R. J. An integrative view of microbiome-host interactions in inflammatory bowel diseases. *Cell Host Microbe* **17**, 577–591 (2015).
2. Imhann, F. et al. Interplay of host genetics and gut microbiota underlying the onset and clinical presentation of inflammatory bowel disease. *Gut* **67**, 108–119 (2018).
3. Huttenhower, C., Kostic, A. D. & Xavier, R. J. Inflammatory bowel disease as a model for translating the microbiome. *Immunity* **40**, 843–854 (2014).
4. Morgan, X. C. et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol.* **13**, R79 (2012).
5. Gevers, D. et al. The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* **15**, 382–392 (2014).
6. Haberman, Y. et al. Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. *J. Clin. Invest.* **124**, 3617–3633 (2014).
7. Lane, E. R., Zisman, T. L. & Suskind, D. L. The microbiota in inflammatory bowel disease: current and therapeutic insights. *J. Inflamm. Res.* **10**, 63–73 (2017).
8. Blander, J. M., Longman, R. S., Iliev, I. D., Sonnenberg, G. F. & Artis, D. Regulation of inflammation by microbiota interactions with the host. *Nat. Immunol.* **18**, 851–860 (2017).
9. Dorrestein, P. C., Mazmanian, S. K. & Knight, R. Finding the missing links among metabolites, microbes, and the host. *Immunity* **40**, 824–832 (2014).
10. McHardy, I. H. et al. Integrative analysis of the microbiome and metabolome of the human intestinal mucosal surface reveals exquisite inter-relationships. *Microbiome* **1**, 17 (2013).
11. Wu, G. D. Diet, the gut microbiome and the metabolome in IBD. *Nestle Nutr. Inst. Workshop Ser.* **79**, 73–82 (2014).
12. Kim, S., Kim, J.-H., Park, B. O. & Kwak, Y. S. Perspectives on the therapeutic potential of short-chain fatty acid receptors. *BMB Rep.* **47**, 173–178 (2014).
13. Smith, P. M. et al. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* **341**, 569–573 (2013).

14. Fernando, M. R., Saxena, A., Reyes, J.-L. & McKay, D. M. Butyrate enhances antibacterial effects while suppressing other features of alternative activation in IL-4-induced macrophages. *Am. J. Physiol. Gastrointest. Liver Physiol.* **310**, G822–G831 (2016).
15. Marchesi, J. R. et al. Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. *J. Proteome Res.* **6**, 546–551 (2007).
16. Wikoff, W. R. et al. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc. Natl Acad. Sci. USA* **106**, 3698–3703 (2009).
17. Williams, B. B. et al. Discovery and characterization of gut microbiota decarboxylases that can produce the neurotransmitter tryptamine. *Cell Host Microbe* **16**, 495–503 (2014).
18. Zelante, T. et al. Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. *Immunity* **39**, 372–385 (2013).
19. Lamas, B. et al. CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. *Nat. Med.* **22**, 598–605 (2016).
20. Le Gall, G. et al. Metabolomics of fecal extracts detects altered metabolic activity of gut microbiota in ulcerative colitis and irritable bowel syndrome. *J. Proteome Res.* **10**, 4208–4218 (2011).
21. Bjerrum, J. T. et al. Metabonomics of human fecal extracts characterize ulcerative colitis, Crohn's disease and healthy individuals. *Metabolomics* **11**, 122–133 (2015).
22. De Preter, V. et al. Faecal metabolite profiling identifies medium-chain fatty acids as discriminating compounds in IBD. *Gut* **64**, 447–458 (2015).
23. Jansson, J. et al. Metabolomics reveals metabolic biomarkers of Crohn's disease. *PLoS ONE* **4**, e6386 (2009).
24. Kolho, K.-L., Pessia, A., Jaakkola, T., de Vos, W. M. & Velagapudi, V. Faecal and serum metabolomics in paediatric inflammatory bowel disease. *J. Crohns Colitis* **11**, 321–334 (2017).
25. Jacobs, J. P. et al. A disease-associated microbial and metabolomics state in relatives of pediatric inflammatory bowel disease patients. *Cell. Mol. Gastroenterol. Hepatol.* **2**, 750–766 (2016).
26. Melnik, A. V. et al. Coupling targeted and untargeted mass spectrometry for metabolome-microbiome-wide association studies of human fecal samples. *Anal. Chem.* **89**, 7549–7559 (2017).
27. Sokol, H. & Seksik, P. The intestinal microbiota in inflammatory bowel diseases: time to connect with the host. *Curr. Opin. Gastroenterol.* **26**, 327–331 (2010).
28. Joossens, M. et al. Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut* **60**, 631–637 (2011).
29. Sokol, H. et al. Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. *Inflamm. Bowel Dis.* **15**, 1183–1189 (2009).
30. Wishart, D. S. et al. HMDB: the Human Metabolome Database. *Nucleic Acids Res.* **35**, D521–D526 (2007).
31. Mosli, M. H. et al. C-reactive protein, fecal calprotectin, and stool lactoferrin for detection of endoscopic activity in symptomatic inflammatory bowel disease patients: a systematic review and meta-analysis. *Am. J. Gastroenterol.* **110**, 802–819 (2015).
32. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
33. Duboc, H. et al. Connecting dysbiosis, bile-acid dysmetabolism and gut inflammation in inflammatory bowel diseases. *Gut* **62**, 531–539 (2013).
34. Abdel Hadi, L., Di Vito, C. & Riboni, L. Fostering inflammatory bowel disease: sphingolipid strategies to join forces. *Mediators Inflamm.* **2016**, 3827684 (2016).
35. An, D. et al. Sphingolipids from a symbiotic microbe regulate homeostasis of host intestinal natural killer T cells. *Cell* **156**, 123–133 (2014).
36. Braun, A. et al. Alterations of phospholipid concentration and species composition of the intestinal mucus barrier in ulcerative colitis: a clue to pathogenesis. *Inflamm. Bowel Dis.* **15**, 1705–1720 (2009).
37. Qi, Y. et al. PPARα-dependent exacerbation of experimental colitis by the hypolipidemic drug fenofibrate. *Am. J. Physiol. Gastrointest. Liver Physiol.* **307**, G564–G573 (2014).
38. Fischbeck, A. et al. Sphingomyelin induces cathepsin D-mediated apoptosis in intestinal epithelial cells and increases inflammation in DSS colitis. *Gut* **60**, 55–65 (2011).
39. Heimerl, S. et al. Alterations in intestinal fatty acid metabolism in inflammatory bowel disease. *Biochim. Biophys. Acta* **1762**, 341–350 (2006).
40. Hove, H. & Mortensen, P. B. Influence of intestinal inflammation (IBD) and small and large bowel length on fecal short-chain fatty acids and lactate. *Dig. Dis. Sci.* **40**, 1372–1380 (1995).
41. Stuart, J. M., Segal, E., Koller, D. & Kim, S. K. A gene-coexpression network for global discovery of conserved genetic modules. *Science* **302**, 249–255 (2003).
42. Wolfe, C. J., Kohane, I. S. & Butte, A. J. Systematic survey reveals general applicability of" guilt-by-association" within gene coexpression networks. *BMC Bioinformatics* **6**, 227 (2005).
43. Frank, D. N. et al. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl Acad. Sci. USA* **104**, 13780–13785 (2007).
44. Lewis, J. D. et al. Inflammation, antibiotics, and diet as environmental stressors of the gut microbiome in pediatric Crohn's disease. *Cell Host Microbe* **18**, 489–500 (2015).
45. Desbois, A. P. & Smith, V. J. Antibacterial free fatty acids: activities, mechanisms of action and biotechnological potential. *Appl. Microbiol. Biotechnol.* **85**, 1629–1642 (2010).
46. German, J. B. & Dillard, C. J. Saturated fats: a perspective from lactation and milk composition. *Lipids* **45**, 915–923 (2010).
47. Galland, L. Magnesium and inflammatory bowel disease. *Magnesium* **7**, 78–83 (1988).
48. Lih-Brody, L. et al. Increased oxidative stress and decreased antioxidant defenses in mucosa of inflammatory bowel disease. *Dig. Dis. Sci.* **41**, 2078–2086 (1996).
49. Yang, J. Y. et al. Molecular networking as a dereplication strategy. *J. Nat. Prod.* **76**, 1686–1699 (2013).
50. Jaskowski, T. D., Litwin, C. M. & Hill, H. R. Analysis of serum antibodies in patients suspected of having inflammatory bowel disease. *Clin. Vaccine Immunol.* **13**, 655–660 (2006).
51. Tigchelaar, E. F. et al. Cohort profile: LifeLines DEEP, a prospective, general population cohort study in the northern Netherlands: study design and baseline characteristics. *BMJ Open* **5**, e006772 (2015).
52. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
53. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
54. Segata, N. et al. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* **9**, 811–814 (2012).
55. Franzosa, E. et al. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat. Methods* **15**, 962–968 (2018).
56. Suzek, B. E. et al. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* **31**, 926–932 (2015).
57. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).
58. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **45632**, D158–D169 (2017).

Acknowledgements

The authors are grateful to the members of the PRISM, LifeLines DEEP and NLIBD cohorts for participating in the study and providing sample material. We thank T. Poon for project management and coordination of data generation, T. Reimels for editorial assistance and A. Garner for providing helpful feedback on the manuscript. The Dutch research team was funded by: CVON IN-CONTROL (CVON2012–03 to A.Z. and J.F.); the Dutch Digestive Foundation (D16–14 to R.K.W. and A.Z.); the Netherlands Organization for Scientific Research (NWO-VIDI 864.13.013 to J.F., NWO-VIDI 016. Vidi.178.056 to A.Z., NWOOW-VIDI 016.136.308 to R.K.W.); a Spinoza Prize (SPI 92–266 to C.W.); and the European Research Council (ERC-Starting no. 715772 to A.Z. and ERC-Advanced 2012–322698 to C.W.). The Boston research team was funded by: the National Science Foundation (NSF CAREER DBI-1053486 and NSF EAGER MCB-1453942 to C.H.); the National Institutes of Health (R01HG00596 to C.H., U54DK102557 to C.H. and R.J.X., R01DK92405 to R.J.X., R24DK110499 to C.H.); the Crohn's and Colitis Foundation of America to R.J.X. and C.H.; and the Center for Microbiome Informatics and Therapeutics (6933665 PO no. 5710004058 to R.J.X.). A.B.H. is a Merck Fellow of the Helen Hay Whitney Foundation.

Author contributions

E.A.F., A.S-M., H.V., C.H. and R.J.X. designed the research. E.A.F., A.S-M., J.A-P., N.F., T.V., H.M. and L.J.M. performed the research. H.J.H., S.R., J.S.S., R.G.W., B.W.S., F.I., A.Z., J.F., R.K.W. and C.W. contributed materials. J.A-P., J.M.S., K.P., A.A.D., K.B. and C.B.C. generated the data. E.A.F., A.S-M., J.A-P., N.F., A.B.H. and H.V. analysed the data. E.A.F., J.A.P., R.K.W., C.W., H.V., C.H. and R.J.X. provided project oversight. E.A.F., A.S-M., H.V., C.H. and R.J.X. wrote the paper.

Competing interests

F.I. received a speaker's fee from AbbVie.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41564-018-0306-4.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to C.H. or R.J.X.

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

nature research

Corresponding author(s): Curtis Huttenhower, Ramnik Xavier

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

No sample-size calculation was performed. Controlling for clinical covariates and correcting for multiple hypothesis testing, we found 100s of strong, statistically significant, and biologically relevant associations between multi'omic features and IBD diagnosis in this cohort. In other words, the study was sufficiently well-powered to produce many useful results.

2. Data exclusions

Describe any data exclusions.

As discussed in the main text, four (4) subjects from the PRISM cohort with metabolomic profiles but NOT metagenomic profiles were excluded from the final analysis. Because the main goal of this work was to compare metagenomic and metabolomic data in the context of IBD, any subjects lacking either of the two data types would have to be excluded. This was a pre-established criterion of the study.

3. Replication

Describe whether the experimental findings were reliably reproduced.

As discussed in the main text, the vast majority of individual trends replicated (at least in sign) within an independent validation cohort. In addition, global properties of subjects' multi'omic profiles were sufficient to predict disease status in the validation cohort with ~80% accuracy.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Not applicable. (This is a cross-sectional study of multiple pre-defined cohorts of individuals subdivided by pre-defined diagnosis.)

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Blinding was not applicable during data collection because study subjects had already been diagnosed in order to be recruited for the study. All samples were de-identified without any diagnostic information prior to data generation. During analysis, diagnosis was always used as a variable in the models, therefore blinding would have been impossible.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

Metabolomic data were pre-processed with Genedata Expressionist v9.0. Metagenomic data were pre-processed with the free and open bioBakery metagenomics workflow (kneadData v0.5.1 for read-level quality control, MetaPhlAn2 v2.2.0 for taxonomic profiling, and HUMAnN2 v0.9.4 for functional profiling). Statistical analyses were carried out with free and open packages in Python (scipy, matplotlib, statsmodels, and scikit-learn) and R (vegan). Additional software details are provided in the main text and Code Availability statement.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

Stool samples used in the study were largely exhausted during metabolomic and metagenomic screening. All chemical reagents, including the library of chemical standards used for metabolome characterization, are readily available through commercial vendors. Specifically, molecules and media reagents were purchased though Sigma-Aldrich, VWR, and ATCC.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used.

b. Describe the method of cell line authentication used.

No eukaryotic cell lines were used.

c. Report whether the cell lines were tested for mycoplasma contamination.

No eukaryotic cell lines were used.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

No commonly misidentified cell lines were used.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	No animals were used.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.	PRISM Cohort: total N=155; diagnosis: 34 non-IBD control, 53 UC, 68 CD; sex: 74 male, 81 female; age: 41.7 ± 16.9 yr (mean ± std. dev.). Netherlands cohorts: total N=65; diagnosis: 22 non-IBD control, 20 CD, 23 UC; sex: 16 male, 27 female, 22 not listed; age: 45.4 ± 15.5 yr.