Gut microbial β-glucuronidases regulate host luminal proteases and are depleted in irritable bowel syndrome

Adam L. Edwinson¹, Lu Yang², Stephanie Peters ³, Nikita Hanning ¹,³, Patricio Jeraldo⁴, Pratik Jagtap ⁵, Joshua B. Simpson⁶, Tzu-Yi Yang ⁵, Praveen Kumar ⁵, Subina Mehta⁵, Asha Nair², Margaret Breen-Lyles¹, Lakshmikanth Chikkamenahalli¹, Rondell P. Graham ⁷, Benedicte De Winter ⁸, Robin Patel ⁵, Surendra Dasari², Purna Kashyap ¹, Timothy Griffin ⁵, Jun Chen², Gianrico Farrugia ¹, Matthew R. Redinbo ⁶, ¹⁰ and Madhusudan Grover ⁵, ⁷, ¹²

Intestinal proteases mediate digestion and immune signalling, while increased gut proteolytic activity disrupts the intestinal barrier and generates visceral hypersensitivity, which is common in irritable bowel syndrome (IBS). However, the mechanisms controlling protease function are unclear. Here we show that members of the gut microbiota suppress intestinal proteolytic activity through production of unconjugated bilirubin. This occurs via microbial β-glucuronidase-mediated conversion of bilirubin conjugates. Metagenomic analysis of faecal samples from patients with post-infection IBS (n = 52) revealed an altered gut microbiota composition, in particular a reduction in Alstipes taxa, and high gut proteolytic activity driven by specific host serine proteases compared with controls. Germ-free mice showed 10-fold higher proteolytic activity compared with conventional mice. Colonization with microbiota samples from high proteolytic activity IBS patients failed to suppress proteolytic activity in germ-free mice, but suppression of proteolytic activity was achieved with colonization using microbiota from healthy donors. High proteolytic activity mice had higher intestinal permeability, a higher relative abundance of Bacteroides and a reduction in β-glucuronidase activity compared with low proteolytic activity mice. High proteolytic activity IBS patients had lower fecal β-glucuronidase activity and end-products of bilirubin deconjugation. Mice treated with unconjugated bilirubin and β-glucuronidase-overexpressing E. coli significantly reduced proteolytic activity, while inhibitors of microbial β-glucuronidases increased proteolytic activity. Together, these data define a disease-relevant mechanism of host–microbial interaction that maintains protease homoeostasis in the gut.

Proteases constitute approximately 2% of the human genome and serve several important biological functions. Infectious, inflammatory, coagulation and neurodegenerative disorders have been attributed in part to aberrant signalling involving the actions of various proteases. The intestinal tract is enriched with a diverse repertoire of proteases of host, microbial and dietary origin. Tight regulation of proteolytic activity (PA) is necessary to prevent tissue injury. In the absence of this regulation, proteases can induce disruption of intestinal barrier function as well as the neuronal sensitization reported in irritable bowel syndrome (IBS). PA in the intestinal tract is poorly understood. We hypothesized that microbiota perturbation induced by a gastrointestinal infection produces ineffective control of PA, allowing the development of PI-IBS.

Gastrointestinal infections are one of the most common events inciting the development of IBS, with the subsequent condition termed post-infection IBS (PI-IBS). Infections promote an immune response and perturb the commensal microbiota. While some data show immune activation in PI-IBS patients, other work found no evidence for these changes. Microbiota data have primarily been descriptive, and certain microbiota changes have been correlated with alterations in pathways associated with epithelial barrier function. No disease-modifying therapies exist for IBS, with treatment largely limited to symptom control. Attempts to restore microbiota using faecal microbiota transplantation (FMT) have produced mixed results. Similarly, probiotic therapies have had limited success possibly due to the absence of a well-characterized, targeted mechanism.

Pancreatic secretions contain large amounts of digestive proteases and these secretions transit through the intestinal tract as chyme. Older studies have demonstrated markedly higher PA in ileostomy contents compared with faeces. Additionally, germ-free (GF) and microbiota-depleted rodents demonstrate high faecal PA. These observations suggest that commensal microbiota directly or indirectly modulate gut PA. Serpins, elafins, siropins and miropins are recognized as inhibitors of specific protease activity; however, the broader endogenous regulation of PA in the intestinal tract is poorly understood. We hypothesized that microbiota perturbation induced by a gastrointestinal infection produces ineffective control of PA, allowing the development of PI-IBS.
One potential pathway for protease inhibition is through unconjugated bilirubin, a potent inhibitor of proteases. Bilirubin is deconjugated in the gut by microbial β-glucuronidase (GUS) enzymes, which have been shown to decrease trypsin and chymotrypsin activity in an in vivo colitis model. The microbiome and the metabolites it creates via enzymes like GUS suggest it may play a larger role in PA regulation than previously understood. We speculated that the healthy microbiota suppresses PA through gut microbial GUS activity, and that this activity would be impaired in patients with PI-IBS.

**Results**

**Faecal proteolytic activity associates with PI-IBS severity.** Patients with *Campylobacter jejuni* infection were prospectively followed for potential development of PI-IBS. As shown in our recent epidemiological study, nearly 21% of individuals who were healthy before the infection fulfilled the Rome III criteria for IBS following infection. We used a fluorescein isothiocyanate (FITC)-casein degradation assay to determine PA in faecal supernatants from PI-IBS patients and healthy volunteers, indicating compositional differences in their microbiota (Fig. 1b). Of the peptides and proteins identified, 12% were assigned as proteases. Within high- and low-PA supernatants, 1,413 and 2,116 unique peptides were identified (see Supplementary Table 1 for description of proteomics database). Forty-two percent of peptides in high-PA and 36% in low-PA faecal samples were of human origin. Only 3 peptides were differentially abundant, all of which were serine proteases of human origin, and all were higher in abundance in high-PA faecal supernatants (Fig. 2b). Chymotrypsin-like pancreatic elastases 2A and 3B both showed 4.5 greater and trypsin 2, a 2.5 greater log-fold abundance in high-PA supernatants. Serine proteases have been shown by us and others to disrupt barrier function. Host colonic mucosa have also been shown to produce serine proteases, including in IBS patients. To determine mucosal expression, bulk RNA sequencing was performed on rectosigmoid colonic mucosal samples from high- and low-PA patients. No differences in expression of proteases or protease inhibitors were observed (see Supplementary Table 2 for complete transcriptomics database).

**High-PA associated microbiota fail to suppress luminal PA.** The differences in microbiota, as well as the identification of host proteases driving PA, led us to examine the role of commensal microbiota in the regulation of host proteases. To model microbial effects, we used germ-free (GF) mice (Swiss-Webster). When compared with specific-pathogen-free mice, GF mice had 10-fold higher faecal PA, indicating that the intestinal microbiome plays a role in PA regulation (Fig. 3a). We next humanized GF mice with microbiota from a random set of high-PA IBS, low-PA IBS patients and healthy volunteers (Fig. 3b). Mice humanized with healthy or low-PA IBS microbiota demonstrated significantly lower PA than mice humanized with high PA (Fig. 3c,d). Humanization with healthy and low-PA IBS microbiota resulted in significant inhibition of PA, whereas high-PA humanization showed no change (Fig. 3e).

In addition to measuring the overall PA using the FITC-casein assay, we assessed enzymatic activity using preferential substrates for a range of serine proteases. We found high activities for chymotrypsin, pancreatic elastase and neutrophil elastase in both human donors and humanized mice (Extended Data Fig. 2a,b). We also assessed the active PA in GF mice and found that activities of trypsin-like, chymotrypsin-like, pancreatic and neutrophilic elastases were similar to those in high-PA humanized mice (Extended Data Fig. 2b). In situ zymography was used to assess colonic mucosal PA, which also showed no differences between high- and low-PA biopsies (Extended Data Fig. 3a,b). Pellet frequency was similar between groups; however, high-PA humanized mice had softer pellets when compared with healthy and low-PA IBS humanized mice (Fig. 3f,g). Colonic length, caecal and mouse weight were similar among the 3 groups (Extended Data Fig. 3c). Small bowel and colon histopathology were evaluated by an expert pathologist in blinded fashion. No gross differences in cellular morphology or distinction in immune cells (polymorphonuclear leukocytes, macrophages) in either the small intestine or the colon were observed between the different humanization states (Extended Data Fig. 3d).

Mice humanized with high-PA faeces have higher intestinal permeability. Next, in vivo permeability assays were performed to investigate potential differences in permeability pathways between mice with distinct faecal humanized states. Differences in pore (<6 Å), leak (<100 Å) and unrestricted (no max) permeability...
Fig. 1 | High PA following C. jejuni infection is characterized by reduced microbiota diversity and taxa loss. a, PA of PI-IBS and healthy volunteers. High PA in 19/52 PI-IBS patients (90th percentile of healthy volunteers, dashed line >1,078 BAEE per mg of protein, n = 38 healthy, 52 PI-IBS). b. PCoA plot of β-diversity. Microbiota composition in high PA is different from healthy (P = 0.05) or low-PA PI-IBS (P = 0.01) (Bray-Curtis distance, PERMANOVA, with 999 permutations, n = 21 healthy, 12 high-PA, 17 low-PA PI-IBS). c. Alpha diversity measures between high-PA PI-IBS, low-PA PI-IBS and healthy volunteers (n = 12 high-PA PI-IBS, 17 low-PA PI-IBS, 21 healthy; linear regressions, *P = 0.05, **P = 0.01, ***P = 0.001). d. Higher-level taxonomic representation (n = 12 high-PA PI-IBS, 21 healthy). e, Correlation matrix between faecal PA and taxa. Correlations made for taxa with r < 0.2 with taxa in red, highlighting differential abundance at q < 0.1. Spearman correlation coefficient (r) identify taxa correlating with PA. Density reflect of proportions of taxa within healthy or high PA, with dots denoting differential at q < 0.1. f, Representative scatterplots between logPA and taxa of entire cohort. High PA and healthy volunteers plotted (n = 21 healthy, 12 high-PA PI-IBS). Grey area: 95% confidence level of linear regression. g, Differentially abundant taxa between healthy and high-PA PI-IBS, Alistipes putredinis absent from high PA (n = 21 healthy, 12 high PA, q < 0.1). Data are presented as mean + s.d. h, Boruta feature selection algorithm identifying taxa predictive of PA status. Six taxa identified as predictive of PA (red), with dark grey defining 3 parameters, shadow max, shadow min and shadow mean produced by Boruta algorithm (n = 12 high PA, 17 low-PA PI-IBS). i, Receiver operating curve of random forests prediction. AUC = 0.813 (95% CI 0.699–0.94). Grey area: 95% confidence interval. Boxplots: lower, middle and upper bounds correspond to the 25th, 50th and 75th percentiles. Upper and lower whiskers extend to largest and smallest value no further than 1.5 x interquartile range (IQR) from respective percentiles.
Fig. 2 | Faecal and tissue proteomics demonstrate that serine proteases of human pancreatic origin drive high PA in PI-IBS. a, Pipeline used for identifying human and microbial faecal proteomic profiles of high-PA and low-PA faeces. b, Metaproteomic analysis of high- and low-PA faecal samples. Volcano plot highlights that faeces from high-PA volunteers have an increased abundance of host pancreatic proteases (n = 7 high PA, 6 low PA, FDR < 0.05). Three proteases—chymotryptsin-like pancreatic elastase 2A, 3B (FDR = 0.002 for each, standard t-test) and trypsin 2 (FDR = 0.02, standard t-test)—all of the serine family, were identified in greater abundance in high-PA samples. c, d, Proteomic analysis of rectosigmoid colonic biopsies using SOMAScan platform. Comparison of proteases (c) and protease inhibitors (d) reveals that production of mucose-derived proteases and inhibitors are comparable regardless of PA phenotype (n = 7 high PA, 6 low PA). e, In vitro inhibition of high-PA FSN with protease inhibitors. AEBSF (100 µM *P = 0.042), nafamostat (100 µM *P = 0.002), UAMC-0050 (100 µM ****P = 0.0001, 10 µM ***P = 0.0002, 1 µM ***P = 0.0004) and elafin (100 µM ***P = 0.0005, 10 µM *P = 0.032) significantly inhibited PA in vitro, while dabigatran (thrombin inhibitor) and E64 (cysteine protease inhibitor) had no effect (two-way ANOVA, Tukey’s multiple comparisons test, n = 3 high- and 3 low-PA FSNs per condition). Barplots are presented as mean ± s.d.
pathways were measured by evaluating serum levels of the orally administered tracer molecules creatinine (6 Å), 4 kDa FITC-dextran (28 Å) and 70 kDa rhodamine B-dextran (120 Å), respectively. Compared with healthy and low-PA PI-IBS humanized mice, high-PA humanized mice had greater serum creatinine and 4 kDa FITC-dextran (Fig. 3h,i). Serum levels of rhodamine B-Dextran were similar across groups, indicating that humanization with high-PA microbiota did not cause increased permeability through the unrestricted pathway (Fig. 3j).

Absence of *A. putredinis* associated with high PA in mice. Mice humanized with either low-PA or high-PA microbiota were taxonomically and compositionally more similar within their given cohort than when compared against each other (Fig. 4a,b). High-PA humanized mice had lower alpha diversity than mice humanized with low-PA or healthy volunteer microbiota (Fig. 4c). This indicates distinct microbiome compositions between high-PA PI-IBS and healthy or low-PA PI-IBS humanized mice (Fig. 4d). Similar to human samples before colonization, no major differences were found at higher taxonomic levels post humanization, with the exception of both Rikenellaceae and Bacteroidaceae which decreased in high-PA mice (Fig. 4e and Extended Data Fig. 4). Differential abundance analysis found 32 taxa at the species level that were differentially abundant between the healthy volunteer and high-PA PI-IBS, while microbiota that were higher in PI-IBS had lower abundance.

**Fig. 3 | Gnotobiotic mice demonstrate that healthy commensal microbiota suppresses host-derived PA while high-PA PI-IBS microbiota does not.**

a. Absence of microbiota characterized by higher PA (Swiss-Webster female, two-tailed Mann-Whitney, n = 5 germ-free, – 3 conventional mice, P = 0.04). b. Mouse model of humanization. c. Healthy and low-PA microbiota causes host faecal PA drop (n = 6 faeces per phenotype, datapoints are mean ± s.d). d. Post-humanization PA is dependent on microbiota. Mice with healthy or low-PA PI-IBS microbiota have significantly lower PA compared with high-PA humanized mice (383 vs 246 vs 1,761 BAEE per mg protein, two-way ANOVA-Tukey’s, n = 6 faeces per phenotype; healthy 150, low PA 13.7, high PA 23.9; healthy 116.3 ± 69.7, one-way ANOVA-Kruskal-Wallis, n = 6 faeces per phenotype, datapoints represent averages, *P* = 0.03 and *P* = 0.04, respectively). e. Low-PA microbiota suppresses host PA, while mice with high-PA microbiota have faecal PA similar to GF mice (% of baseline, healthy 24.5 ± 32.8, low PA 17.8 ± 23.9; high PA 116.3 ± 109.7, one-way ANOVA-Kruskal-Wallis, n = 6 faeces per phenotype, datapoints represent averages, *P* = 0.03). f. Humanized mouse pellet frequency. No difference between humanization states in pellet production. g. Humanized mouse faecal pellet consistency. High-PA humanized mice have significantly looser faeces. Scored 0, normal to 4, diarrhoea. Two averaged independent observations – 6 faeces per phenotype, datapoints represent averages, *P* = 0.03. h. In vivo permeability. Absence of the unrestricted pathway (Fig. 3j). i. Leak pathway permeability. The 4 kDa FITC-dextran levels in high-PA humanized mice show increased leak pathway permeability (high-PA PI-IBS, 19.1 ± 14.6 mg dl⁻¹; low-PA PI-IBS 23.9 ± 23.9; healthy 13.7 ± 30.3, one-way ANOVA-Kruskal-Wallis, *P* = 0.04). j. Unrestricted pathway permeability. No differences in serum rhodamine levels across humanization states. Boxplots in d–j as described in Fig. 1.
Fig. 4 | Microbial diversity and composition in humanized mice identify specific microbial taxa predictive of PA status and metabolic pathway differences. a, Mice humanized with high-PA, low-PA PI-IBS and healthy volunteer microbiota have differences in microbiota composition (n = 6 human faeces per phenotype, datapoints represent averages). b, Humanized mice intra-group, not inter-group relatedness. Humanized mice microbiota are compositionally and taxonomically different when compared to each other, but similar within its group (Bray-Curtis, PERMANOVA with 999 permutations, P = 0.001). c, Alpha diversity across humanized mice. Healthy volunteer and low-PA humanized mice have greater species richness compared with high-PA humanized mice (n = 6 human faeces per phenotype, dots represent averages, linear regression on observed species, **P = 0.01, ***P = 0.002). d, Beta diversity measures in humanized mice. Microbial compositions differ between humanization states (Bray-Curtis, PERMANOVA with 999 permutations, P = 0.01). e, Higher-level taxonomic evaluation of engrafted microbiota in humanized mice (n = 6 human faeces per phenotype). f,g, Volcano plots highlighting strain-level differences between high-PA and healthy volunteer (f) or high- and low-PA PI-IBS humanized mice (g). Red datapoints indicate differences in abundance (q < 0.1, dotted line), with A. putredinis identified (n = 6 human faeces per phenotype, datapoints represent averages). h, Prediction model of microbiota for PA status by random forests. The top 10 taxa predictive of PA status in mice, assessed using mean decrease in accuracy. i, Predicted KEGG pathway differences between high-PA and healthy volunteer humanized mice. Differentially abundant KEGG pathways with q < 0.1. Effect size denotes average difference in KEGG functional unit between reference and comparison groups. Boxplots as previously described.
identified the genus *Alistipes* as a major predictor of low-PA status, with *A. putredinis* being the most significantly associated species (Fig. 4h). Furthermore, an AUC ROC of 0.914 was generated (Extended Data Fig. 5c). Nine predicted Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were predicted to be significantly different between the healthy and high-PA groups (Extended Data Fig. 5d). The top predicted pathways elevated in high-PA humanized mice were ‘aromatic degradation’ and ‘mineral and organic ion transport’, whereas the ‘polyamine biosynthesis’ pathways were increased in healthy humanized mice (Fig. 4i).

*Faecal microbiota restoration and *A. putredinis* supplementation restores protease suppression.* Introducing microbial taxa associated with low PA into the microbiome of high-PA individuals may serve as a potential way to lower PA. One strategy for introducing new taxa is through FMT using a healthy low-PA donor. High-PA PI-IBS humanized mice were administered either a vehicle control or a faecal slurry from a healthy, low-PA donor (1.62 No-benzoyl-L-arginine ethyl ester solution (BAEE) units per mg protein). Faecal PA of FMT-treated mice was significantly lower than that of untreated control mice (Fig. 5a) and significant inhibition of PA from baseline was observed in FMT-treated but not in control mice. On measurement of specific activity, a reduction in the activity of all serine proteases, except chymotrypsin and kallikrein (KLK), was observed in mice given an FMT compared with the control group (Fig. 5b).

Mice given an FMT had greater alpha diversity (Fig. 5c) as well as beta diversity (Fig. 5d) than vehicle controls, indicating greater species diversity and changes in microbiota composition post FMT. While no differences were observed at phyla and class taxonomic levels, Rikenellaceae, the family of *A. putredinis*, and 8 additional families were differential post FMT (Extended Data Fig. 6a). At the species level, there were 35 differentially abundant taxa between the two groups post FMT (Fig. 5e). Shifts in KEGG pathways post FMT included increased ‘methylene metabolism’ and decreased ‘aromatics degradation’ (Fig. 5f, q < 0.1). Interestingly, metabolic by-products of aromatic degradation include amino acid derivatives like phenylethylamine, which are increased in Crohn’s disease30. Of the differentially abundant taxa, 13 were in greater abundance in the mice receiving FMT. Multiple species of the genus *Alistipes* were identified (Fig. 5g), which correlated with decreased faecal PA (Extended Data Fig. 6b).

In humans and humanized mice, the presence of *A. putredinis* in the gut microbiota strongly predicted low-PA status. We aimed to test if *A. putredinis* could reverse high-PA phenotype by: (1) giving high-PA humanized mice an FMT with a community containing *A. putredinis* (A.p.+) or with a community lacking *A. putredinis* (A.p.−); and (2) adding *A. putredinis* directly to a high-PA faecal slurry used to humanize mice. Mice given an FMT with an *A. putredinis*+ healthy community had a robust lowering of PA when compared with mice receiving an FMT with an *A. putredinis*− healthy donor community (Fig. 5h). The addition of *A. putredinis* to a high-PA faecal slurry also attenuated the marginal increase that was seen with the high-PA humanization (Fig. 5i).

**GUS-mediated production of unconjugated bilirubin suppresses PA.** The lack of a known protease inhibitor peptide led us to investigate an alternative mechanism by which luminal proteases could be regulated by the microbiome. Previous literature has correlated the presence of *A. putredinis* with faecal beta-glucuronidase (GUS) enzyme activity31,32, and *Alistipes* GUS enzymes have been identified in the human and mouse gut microbiomes31,33. Therefore, we investigated the faecal GUS activity of high-PA PI-IBS individuals. High-PA PI-IBS patients had significantly lower faecal GUS activity compared with healthy or low-PA PI-IBS individuals (Fig. 6a). We then employed an untargeted high-performance liquid chromatography mass spectrometry (HPLC–MS/MS) metabolomics approach to examine microbial GUS-related metabolites. While a number of differentially abundant metabolites were observed, high-PA samples had significantly lower levels of urobilinogen (Fig. 6b,c, q < 0.05). Urobilinogen is generated only when conjugated bilirubin is first deconjugated by gut microbial GUS enzymes.

Increased GUS activity and urobilinogen levels in low-PA individuals led us to test whether the administration of GUS could inhibit PA in mice. We monoclonized GF mice (which started with high PA) with either competent control *Escherichia coli*, or a modified strain of *E. coli* engineered to constitutively overexpress GUS (GUS⁎). Mice given the GUS⁎ *E. coli* had significantly lower PA than mice receiving a vehicle control gavage (Fig. 6d). Bilirubin mono- and di-glucuronides (conjugated bilirubin) are key endogenous substrates for gut microbial GUS enzymes. Unconjugated bilirubin has been shown to inhibit serine proteases in vitro34. We tested the in vitro capacity of various metabolites in the bilirubin deconjugation pathway to inhibit purified trypsin. Only unconjugated bilirubin was able to suppress the PA significantly. Mesobilirubin and urobilinogen had some effect, but this was not statistically significant and conjugated bilirubin had no effect (Extended Data Fig. 7a,b). Enzyme kinetics and inhibitory concentrations for the tested bilirubin metabolites are provided in Extended Data Table 2.

Next, we determined the effects of unconjugated bilirubin on PA in vivo. We administered unconjugated bilirubin (200 μM) orally for 3 d in GF mice. This resulted in a significant reduction in trypsin-like activity when compared with untreated control mice (Fig. 6e). We also found that n-urobilin, an end product of bilirubin deconjugation, was increased in GUS⁎ *E. coli*-treated mice (ratio of control/treated = −12.3, P = 0.001) (Fig. 6f). None of the other identified products of bilirubin metabolism pathways showed statistically significant differences. Taken together, these data support the conclusion that both GUS and unconjugated bilirubin play important roles in reducing gut luminal PA.

**Inhibition of gut microbial GUS enzymes increases host PA.** We next inhibited gut microbial GUS activity in healthy humanized mice to determine effects on the PA (Fig. 6g). Mice receiving d-glucaric-1,4-lactone, a broad-spectrum but modestly potent inhibitor of GUS activity35, demonstrated significantly higher PA than mice given a control gavage (Fig. 6h). These mice had decreased levels of bilirubin degradation product, G17H2N20S5 (ratio of treated/control = −10, P = 0.03; Fig. 6i) and increased permeability for the 4 kDa FITC-dextran (Extended Data Fig. 7c). We next employed UNC10201652, a more potent GUS inhibitor that exclusively inhibits gut microbial GUS enzymes36. Using a similar dosing strategy (Fig. 6g), healthy humanized mice given UNC10201652 demonstrated significantly higher PA compared with control animals (Fig. 6i). Thus, the robust increases in PA produced by GUS inhibitors, and specifically by UNC10201652, indicates that the microbiome is critical for regulation of host PA and microbial GUS activity is clearly important for reducing luminal PA in vivo (Extended Data Fig. 7d).

**Discussion** Imbalance in intestinal proteolytic activity has been demonstrated in various conditions, including IBS32,33,37,42, IBD43,44,47 and coeliac disease.48-50 However, the regulation of PA in the intestinal tract is not well understood. Here we demonstrate that specific proteases of host origin drive PA and that commensal microbes inhibit proteases through gut microbial β-glucuronidase-mediated production of unconjugated bilirubin. A characteristic feature of the high-PA phenotype is the absence of commensal *A. putredinis*, and introduction of *A. putredinis* containing low-PA community into high-PA mice or supplementing with *A. putredinis* alone can suppress the PA. These observations support the conclusions that PI-IBS...
Fig. 5 | Faecal microbiome transfer of low-PA microbial communities lowers PA of high-PA humanized mice in an Alistipes-dependent manner. 

a. FMT treatment in high-PA mice. FMT with a healthy microbiome decreased PA in high-PA humanized mice compared with controls (1,676 ± 613.6 vs 3,920 ± 2,178 BAEE units per mg protein, one-way ANOVA-Kruskal-Wallis *P = 0.0235 and two-sided Wilcoxon-matched pairs *P = 0.011, 9 mice per group). 

b. FMT changes proteolytic profile. Trypsin, neutrophilic elastase and pancreatic elastase activities were lower post FMT compared with baseline (paired t-test, Mann-Whitney, n = 5 mice, trypsin *P = 0.003, pancreatic elastase *P = 0.04, neutrophil elastase *P = 0.006).

c. Measurement of alpha diversity. Mice given FMT demonstrate greater microbial diversity compared with control mice (linear regression on observed species, n = 9 per group, P = 0.01).

d. Measurement of beta diversity. Bray-Curtis β-diversity ordination demonstrates shifts in microbiota composition following FMT. Mice receiving FMT cluster separately and towards a healthy, low-PA microbiome (PERMANOVA, n = 9 mice per group, P = 0.001).

e. Heat map outlining differentially abundant taxa in mice treated with FMT. Taxa identified as differentially abundant at q < 0.1 (n = 9 mice per group). 

f. Proportion of differentially abundant bacteria between FMT and control mice. Thirty-five differentially abundant bacterial taxa were identified at the species level (q < 0.1), with 13 having greater abundance after FMT (n = 9 mice per group).

g. Predicted KEGG pathway differences. Methane metabolism pathway increased, and aromatics degradation decreased in FMT mice (n = 9 mice, q < 0.1).

h. Restoration of low-PA phenotype is dependent on FMT community. A list of differentially abundant bacterial taxa in FMT mice compared with controls (n = 9 mice per group).

i. High-PA faeces spiked with A. putredinis attenuates increased PA post humanization (two-way ANOVA-Sidak’s, n = 3 mice per group, P = 0.05, mean ± s.d.). Boxplots as previously described.
Articles Nature Microbiology

12.2 nmol min⁻¹, Unconjugated bilirubin suppresses PA. Mice given unconjugated bilirubin (gavage, 3 d) had less trypsin-like activity compared with controls (4.8 vs 411.7 vs 3,371 = n lower faecal PA compared with controls (1,437 = t-test).

P = < P

-Box plots as previously described. Human and mouse metabolite intensities were normalized using total ion current for each sample.

Beta-glucuronidase activity (BAEE units per mg protein, Mann-Whitney, 2,430 vs 113.8 = ± 230.1 BAEE units per mg protein, Mann-Whitney, n = 5, P = 0.008). Data presented as mean ± s.d. Mice given unconjugated bilirubin (gavage, 3 d) had less trypsin-like activity compared with controls (4.8 vs 411.7 vs 3,371 = n lower faecal PA compared with controls (1,437 = t-test).

Human and mouse metabolite intensities were normalized using total ion current for each sample, as mean ± s.d. Boxplots as previously described. Human and mouse metabolite intensities were normalized using total ion current for each sample. High-PA PI-IBS had lower faecal GUS activity compared with healthy and low-PA PI-IBS (one-way ANOVA-Kruskal-Wallis, * = P = 0.008).

Microbiota mediates PA suppression through microbial GUS enzymatic activity and production of unconjugated bilirubin. Mice given unconjugated bilirubin (gavage, 3 d) had less trypsin-like activity compared with controls (4.8 vs 411.7 vs 3,371 = n lower faecal PA compared with controls (1,437 = t-test).

Microbiota mediates PA suppression through microbial GUS enzymatic activity and production of unconjugated bilirubin.

-Oral administration of d-glucaro-1,4-lactone, a non-specific GUS inhibitor, has significantly lower faecal PA compared with controls (1,437 = 411.7 vs 3,371 = 1,437 = 2,430 vs 113.8 = t-test).

8.0 BAEE units per mg protein, Mann-Whitney, 2,430 vs 113.8 = ± 230.1 BAEE units per mg protein, Mann-Whitney, n = 5, P = 0.008). Data presented as mean ± s.d. Boxplots as previously described. Human and mouse metabolite intensities were normalized using total ion current for each sample, as mean ± s.d. Boxplots as previously described. Human and mouse metabolite intensities were normalized using total ion current for each sample, as mean ± s.d. Boxplots as previously described. Human and mouse metabolite intensities were normalized using total ion current for each sample. High-PA PI-IBS had lower faecal GUS activity compared with healthy and low-PA PI-IBS (one-way ANOVA-Kruskal-Wallis, * = P = 0.008).

Fig. 6 | Microbiota mediates PA suppression through microbial GUS enzymatic activity and production of unconjugated bilirubin. a. GUS activity in humans. High-PA PI-IBS had lower faecal GUS activity compared with healthy and low-PA PI-IBS (one-way ANOVA-Kruskal-Wallis, * = P = 0.02 and * = P = 0.03, respectively, n = 4 low PA, 8 high PA and 8 healthy). b. Concentration of urobilinogen in human faeces. Untargeted faecal metabolomics identified higher urobilinogen in low-PA compared with high-PA faeces (n = 3 high PA and 4 low PA).

Protoporphyrin (1)**, Bilirubin degradation product, C17H18N2O4 (2)**, Bilirubin degradation product, C17H18N2O4 (1)**, Bilirubin degradation product, C17H18N2O4 (3)**, Bilirubin degradation product, C17H18N2O4 (2)**, Bilirubin degradation product, C17H18N2O4 (1)**, Bilirubin degradation product, C17H18N2O4 (3)**, Bilirubin degradation product, C17H18N2O4 (1)**, Bilirubin degradation product, C17H18N2O4 (2)/**, Bilirubin degradation product, C17H18N2O4 (3)/**, Bilirubin degradation product, C17H18N2O4 (1)/**, Bilirubin degradation product, C17H18N2O4 (2)/**, Bilirubin degradation product, C17H18N2O4 (3)/**, Bilirubin degradation product, C17H18N2O4 (1)/**, Bilirubin degradation product, C17H18N2O4 (2)/**, Bilirubin degradation product, C17H18N2O4 (3)/**, Bilirubin degradation product, C17H18N2O4 (1)/**, Bilirubin degradation product, C17H18N2O4 (2)/**, Bilirubin degradation product, C17H18N2O4 (3)/**, Bilirubin degradation product, C17H18N2O4 (1)/**, Bilirubin degradation product, C17H18N2O4 (2)/**, Bilirubin degradation product, C17H18N2O4 (3)/**, Bilirubin degradation product, C17H18N2O4 (1)**...
pathophysiology is driven by loss of critical microbes following intestinal infection, which result in an impaired bilirubin deconjugation and weakened inhibition of specific serine proteases.

Due to the complexity of the luminal environment in the GI tract, it has been difficult to assign the nature or origin of proteases that become elevated in disease states. While high- and low-PA patients have microbiota differences, faecal metaproteomics did not identify any major differences in the abundance of bacterial proteases. Previous work has in fact shown that human-derived proteases make up the bulk of faecal proteases. The intestinal epithelium has been identified as a source of luminal proteases, specifically trypsin-3. Furthermore, IBD patients have been found to have elevated tissue elastase and thrombin activity. Here we found no transcriptomic or proteomic signatures of activated immune pathways, making it probable that intraluminal proteases are not driven by immune cells. Further, mucosal PA was similar between high- and low-PA colonic biopsies.

Because intestinal infection and subsequent antibiotic use probably deplete microbiota diversity, we speculated that this feature plays an important role in PA regulation. The presence of A. putredinis was the strongest predictor, with high-PA patients demonstrating an absence of this microbe compared with its 3–5% prevalence in healthy and low-PA PI-IBS patients. Alistipes is a genus of Gram-negative, rod-shaped, non-spoore-forming anaerobic bacteria that currently consists of 13 species. Studies have shown the positive impact of Alistipes spp. as oral administration of A. fungoides decreased colitis in a dextran sulfate sodium (DSS)-mouse model and Alistipes enrichment was associated with anti-inflammatory cytokine production in 2,4,6-trinitrobenzene sulfonic acid (TNBS) colitis in nucleotide-binding oligomerization domain-2 (NOD2) knockout mice. In contrast, one study showed that the development of artificial sweetener-induced ileitis was associated with an increase in Alistipes abundance. Further studies are needed to clarify the roles Alistipes spp. play in microbiome dynamics and intestinal health.

Here we employed GF mice to model the mechanistic influences of the microbiota on host-driven PA. Similar to older observations in GF rats, we found GF mice to have 10-fold higher faecal PA compared with specific-pathogen-free mice. Humanization with healthy commensal microbiota significantly decreased the PA, which was not observed upon humanization with high-PA microbiota. Additionally, high-PA humanized mice had greater permeability through pore and leak pathways but not unrestricted pathways. This and the absence of gross inflammatory changes in the intestine suggest that the aberrant regulation of PA does not induce frank inflammation, similar to what is noted clinically in IBS patients. One earlier study has correlated faecal PA with microbiota composition, and increased PA was also associated with increased permeability. Together with these previous observations, the current study demonstrates an essential role of microbiota-based inhibition of the PA, which protects the colonic epithelium from barrier-disrupting proteases.

While microbiota restoration has been of interest for treating IBS, mixed results have been observed with probiotics and FMT trials have largely been unsuccessful, with a few notable exceptions. In our mouse model, introduction of a healthy low-PA community into high-PA humanized mice via FMT caused a drop in PA and changes in both alpha and beta diversity. Interestingly, we also found that reduced PA was not a universal phenomenon after FMT with healthy donor communities, akin to the variability observed in clinical FMT studies. The presence of low-PA predictors, such as A. putredinis, was critical in lowering PA, as FMT from donors lacking these taxa did not result in suppression of PA. Recent studies have suggested that FMT engraftment depends on both the donor and recipient microbiome pre-FMT, and that successful treatment may be dependent on the presence of specific bacterial taxa. Additional studies are necessary to unravel the essential features of engraftment dynamics, as well as the key microbial taxa that associate with PA suppression and correction of barrier dysfunction.

Inhibition of host proteases could arise from microbial protease inhibitors, microbial influences on host protease inhibitors, and through changes in the luminal milieu, which is critical for the deconjugation of bilirubin. Alistipes β-glucuronidases have also been established to be present in both human and mouse gut microbiomes. Compared with the low-PA PI-IBS individuals, high-PA patients had both lower faecal levels of GUS activity and lower faecal levels of urobilinogen, an end product of deconjugated bilirubin metabolism. We demonstrate that unconjugated bilirubin, but not some of the other products in the bilirubin deconjugation pathway, decrease trypsin activity in vitro and in vivo. Furthermore, we show that E. coli with constitutive GUS expression suppresses PA in vivo and inhibition of GUS converts healthy low-PA humanized mice into high-PA mice. The results we obtained identify a microbiome-dependent mechanism for host PA regulation and also provide an exciting potential therapeutic target for the treatment of the high PA associated with gut diseases.

Although specific taxa such as A. putredinis predict low PA, further studies are required to determine whether their presence is essential for suppression of PA. GUS enzymes are commonly expressed by bacteria and there is marked variability in the efficiency of different gut microbial GUS enzyme orthologues in processing substrates such as oestrogen. As bilirubin processing by microbial GUS enzymes has not been studied in detail, further studies will need to determine whether certain taxa are more efficient at inhibition of PA via this deconjugation activity. It is also plausible that additional named or unnamed metabolites in the bilirubin deconjugation pathway or metabolites other than those in the bilirubin deconjugation pathway also play a role in inhibition of serine proteases.

The tissue and faecal proteomics assays are limited by sensitivity and may have missed low-abundance proteases or protease inhibitors. Determination of specific proteases driving the PA is limited by the poor specificity of the tracers as well as inhibitors. However, the complementary experiments using proteomics, substrate assays and inhibitors all point towards trypsin and elastases as drivers of PA in these patients. Although humanized models have limitations in terms of immune development and the potential for incomplete engraftment of the donor microbiome, our model was found to be consistent with others and was able to recapitulate the PA phenotypes, proteolytic profiles and the physiological changes in permeability observed in humans. The complementary human and mice findings presented here establish that the intestinal microbiome is critical for inhibiting host-derived proteases and maintaining barrier function through their GUS-mediated production of unconjugated bilirubin. They suggest that microbiota-based therapies may potentially target protease imbalances, as well as the resulting phenotypes of impaired barrier function and visceral hypersensitivity in IBS and other gastrointestinal disorders.

Methods

Study design. The goal of this study was to determine the source and regulation of proteases within the gastrointestinal tract of individuals with high faecal proteolytic activity (PA). Within 7 to 10d of notification of Campylobacter cases, the Minnesota Department of Health conducts interviews with patients or guardians who are culture positive. PI-IBS development was initially assessed by sending surveys to patients aged 18–80 years from 6 to 9 months following a positive culture for Campylobacter species, recruiting individuals from 1 November 2011 to 30 September 2019. This was done by mail or by telephone interviews for non-responders. Participants were then asked to complete a Rome III IBS questionnaire. The identified patients were then invited for recruitment at Mayo Clinic. For the study physicians measured IBS and classified that exclusion and inclusion criteria have been met, and participants understand the instructions for collection of stool samples and metadata (bowel diary, psychological and other questionnaires). Healthy volunteers were recruited through advertisements.
They completed a screening questionnaire to confirm absence of ongoing GI symptoms or previous GI diagnoses. Both male and female participants were recruited for the study.

A total of 52 PI-IBS patients defined by Rome III criteria and 38 healthy volunteers were recruited in an observational study. The Rome III criteria were used to define IBS and its subtypes on the basis of the presence of abdominal pain or discomfort at least 2–3 times a month and alterations in bowel function (stool consistency or frequency) for ≥6 months to demonstrate chronicity of symptoms.

Those with a history of abdominal surgery (except hernia, C-section, hysterectomy, appendectomy or cholecystectomy), inflammatory bowel disease, microscopic colitis, or coeliac disease were excluded. Additionally, recruited volunteers were not pregnant at the time of the study. Use of tobacco or alcohol for the duration of the study was prohibited. The following medications were prohibited 7 days before study participation: those affecting gastrointestinal transit, serotoninergic agents, anti-cholinergic agents, antihistamines, antacids, peppermint oil, narcotics, anticonvulsants, or new probiotics. Ingestion of artificial sweeteners such as Splenda (sucralose), Nutrasweet (aspartame), lactulose or mannitol was prohibited for 24 h before the start and during the study. All participants taking part in the study were asked to complete the Hospital Anxiety and Depression Scale (HADS) and a 7 d bowel diary. All participants were issued with the validated IBS-SSS63 and IBS-QoL scales were used. Established Bristol Stool scale (IBS-QoL) questionnaire as well as the Long Bowel Disease questionnaire (BDQ). All participants completed the HADS. PI-IBS patients also completed the Symptom Checklist-90 (SCL-90), IBS Symptom severity score (IBS-SSS), IBS-quality of life (IBS-Qol) questionnaire as well as the Long Bowel Disease questionnaire (BDQ).

Validated IBS-SSS63 and IBS-Qol scales were used. Established Bristol Stool Scale (IBS-QoL) questionnaire as well as the Long Bowel Disease questionnaire. Both male and female participants were recruited for the study. All participants taking part in the study were asked to complete the Hospital Anxiety and Depression Scale (HADS) and a 7 d bowel diary. All participants were issued with the validated IBS-SSS63 and IBS-QoL scales were used. Established Bristol Stool scale (IBS-QoL) questionnaire as well as the Long Bowel Disease questionnaire (BDQ). All participants completed the HADS. PI-IBS patients also completed the Symptom Checklist-90 (SCL-90), IBS Symptom severity score (IBS-SSS), IBS-quality of life (IBS-Qol) questionnaire as well as the Long Bowel Disease questionnaire (BDQ).

measurement kinetically at 37 °C for 15 min. Enzymatic activity was calculated for each enzyme (nmol min−1) using a standard curve generated for free AMC, which was then converted to specific activity (nmol min−1 μg−1) by using the protein concentration obtained using the Bradford method.

In vitro inhibition of high-PA FSNS using protease inhibitors. AEBSF (Sigma Aldrich, A8456), nafamostat mesilate (Sigma Aldrich, N0289), eflafin (Sigma Aldrich, E7280), dabigatran (Sigma Aldrich, SML2370), E64 (Sigma Aldrich, E3132) and UAMC-0050 inhibitors were used to inhibit the proteases present in high-PA and low-PA FSNS. A 5 μl sample of FSNS prepared as previously described, was added to a 96-well black-bottom plate containing 20 μl of the tested inhibitor. Each inhibitor was tested in a final concentration of 10, 1 and 0.1 μM in the wells. FITC-labelled cassein (175 μl) was then immediately added to the wells and allowed to incubate for 10 min at room temperature. Fluorescent intensity as a result of proteolytic digestion of FITC-cassein was read. A standard curve was used to determine proteolytic activity, which was then normalized to the protein content. The sample that had been previously measured by Bradford assay, similar to the aforementioned measurement of faecal proteolytic activity.

In situ zymography. High- and low-PA distal colonic tissues were assessed for potential tissue contributions to PA. Biopsy sections (8 μm) were freshly embedded in optimal cutting temperature compound, air-dried on a glass slide and stained with a green-fluorescent SYTOXX Green nuclear stain (Thermo Fisher, S7020) for 15 min in a humid chamber at room temperature. The sections were then covered with a low-melt agar overlay (0.3%) containing N-2-3-Osyl-Gly-Pro-Pro-Arg-7-amido-4-methylcoumarin hydrochloride (100 μg ml−1) prepared in 50 mM Tris-HCl pH 8.0 containing 10 mM CaCl2. The substrate agar overlay allows for the visualization of localized tissue-specific proteolytic activity.

A single slide for each sample was then incubated at 4°C overnight and served as a background control for that sample, while paired slides were incubated at 37°C overnight and served as the test slides. Fluorescence intensity was measured under the same settings for both the 4°C and 37°C incubated slides, with images obtained using an epifluorescence microscope. Area-corrected image intensities were collected using ImageJ Java 1.8.0, 172 software and reported as the difference between the testing slide and the background control. Representative images were pseudocoloured blue for nuclei and green for tissue enzymatic activity.

Shotgun metagenomics of human samples. Sequencing. DNA was purified from human stool samples using MO BIO Laboratories PowerSoil DNA isolation kit, MP Biomedicals FastPrep-24 5G homogenizer and an automated high-throughput nucleic acid isolation system (Perkin Elmer Chemagen MSN II) based on proprietary M-PVA bead technology. The concentration was measured using Qubit double-stranded, high sensitivity dsDNA HS kit (PN Q32851 Thermo Fisher).

Libraries were sequenced on an Illumina HiSeq 4000, PE150. Sequence reads were performed post-processing and taxonomy calling. Raw reads were processed using BBduk v38.73 to remove adapters and perform read quality filtering.

Bioinformatic. Processed reads were analysed with the SHOGUN v1.0.8 taxonomy profiler, using the BURST aligner option. Profiling of the functional content was performed using SHOGUN v1.0.8 to obtain KEGG orthology entries, modules and pathways, as well as HUMAnN2 v2.8.1 using its UniRef90 database to obtain UniRef90 gene family abundance. Functional profiling against targeted databases was performed with HUMAnN2, using the MEROPS peptidase database release 12.0.1.

Measurement of GUS activity. GUS enzymatic activity was measured using the substrate 4-Nitrophenyl-β-D-glucuronide (10 mM, pH 7.0). FSNS (50 μl) were dispensed in duplicate into a 96-well clear-bottom plate. A plate dispenser was used to submucosally add 50 μl of normal saline to each well. The reaction was moved immediately to a BioTek Synergy MX microplate reader, pre-warmed to 37°C. Plates were read at 405 nm every 5 min for 1 h. Concentrations of GUS (U mg−1 protein) were determined on the basis of a standard curve of pure GUS enzyme and standardized to the total amount of protein determined as the Bradford assay.

Faecal metaproteomics. In-gel digestion. Gel bands were washed using 1:100 mM ammonium bicarbonate/acetonitrile and subsequently dehydrated using acetonitrile. Proteins were reduced and alkylated by using 10 mM dithiothreitol in 50 mM ammonium bicarbonate and 55 mM iodoacetamide in 50 mM ammonium bicarbonate in the dark, respectively. The gel pieces were washed with 1:1 acetonitrile-100 mM ammonium bicarbonate and subsequently dehydrated with acetonitrile. Trypsin (5 μg ml−1 in 50 mM ammonium bicarbonate and 5 mM calcium chloride) was added for in-gel digestion overnight at 37°C. Peptides were extracted sequentially using 50% acetonitrile 0.3% formic acid and 80% acetonitrile 0.3% formic acid. Peptides were desalted using the MCX-type STAGE tips (3 Empore SDB-XC filters) and dried in vacuo for storage at −80°C.

Liquid chromatography. An Eksigent 425 system was coupled to an AB Sciex TripleTOF 5600+ mass spectrometer. The liquid chromatography system uses a trap column (ChromXP 5 μm ChromXP C18CL 120 Å 10 x 0.3 mm) and an analytical column (Eksigent 3 μm ChromXP C18CL 120 Å 150 x 0.3 mm). The analytical column temperature was maintained at 35°C. LCMS grade solvents were purchased from Honeywell and composed of buffer A (0.1% formic acid in water (v/v)) and buffer B (0.1% formic acid in acetonitrile (v/v)). Peptides were reconstituted in 10 μl 2% solvent B and 8 μl was loaded at 10 μl min−1 for 3 min onto the trap column. Analytical separation was established by maintaining 3%
buffer B during loading. After a valve switch event, peptides were resolved using the analytical column at a flow rate of 5 μl min⁻¹, with a linear gradient of 3–25% HPLC buffer B for 38 min, followed by a 5 min linear gradient to 32% buffer B. Following the peptide elution window, in 2 min the gradient was increased to 80% buffer B for 3 min. Initial chromatographic conditions were restored in 1 min and maintained for 8 min.

Mass spectrometry. Peptides were ionized using the SCiEX DuoSpray ion source with a 50 μm ESI electrode. Data were acquired using an ion spray voltage floating at 5.5 kV, curtain gas of 35 psig, GS1 of 19 psig, GS2 of 15 psig and an interface heater temperature of 75 °C. The Information Dependent Acquisition (IDA) mode was used to scan ions with charge states +2 to +5 with intensity greater than 150 units. Exclusion time was 15 s and tolerance was set to 100 ppm. Survey scans (MS1) were acquired in 250 ms over a mass range of 400–1,250 m/z. Following the survey scan, 200 scans were acquired over an mass range of 100–1,500 m/z using a collision energy setting of 35 spreading 5 eV.

Colonic tissue proteomics. Biopsies were collected from the sigmoid colon and completed by a single endoscopist (M.G.). A flexible sigmoidoscopy was done following a tap water enema with biopsies using large biopsy forceps (no pin) 25 cm from the anal verge, and samples were immediately snap frozen in liquid nitrogen and stored at −80 °C until use. These were cryopulverized using a freezer mill. Extraction buffer (200 μl, T-PER tissue protein extraction agent, Thermo Fisher) plus 2 μl of Halt protease inhibitor cocktail (Thermo Fisher) were added to each tube. Samples were homogenized on ice using an Omni PowerGen homogenizer with Omni tips for soft tissues (Thermo Fisher) for 30 s until no fragments were visible. Samples were centrifuged at 14,000 g at 4 °C for 10 min. Supernatants were collected, filtered using a 0.2 μm Spin-X tube and centrifuged at maximum speed at 4 °C for 5 min. Samples were shipped on ice to Washington University St. Louis to be run on the SomaScan 1.3 K cell and tissue platform.

Colonic tissue transcriptomics. Sigmoid colon biopsies were collected from volunteers, placed in RNA later (Fisher) and immediately snap frozen in liquid nitrogen. Total RNA was extracted using a miRNeasy kit from Qiagen following the manufacturer’s instructions and measured with a RiboGreen RNA quantification kit (Invitrogen). RNA integrity and profile were analysed using an Agilent BioAnalyzer/TapeStation, followed by library preparation with Illumina’s TruSeq Stranded mRNA Library preparation kit without polyA selection. Final libraries were measured using Picogreen and quality assessed using the Agilent Tapestation. TruSeq stranded mRNA libraries with unique dual-indexed (UDI) tags were generated (Illumina) for the purpose of sequencing. All libraries were pooled and sequenced on a NovaSeq V2 150bp Charter-Service RNA-seq run, generating ≥2 million reads for each library, with mean quality scores for all libraries ≥Q30 on a 150 PE flow cell lane. RNA-seq fastq files for samples were obtained via an Illumina sequencing platform, with quality control performed on the FASTQ files using RSeQC software version v3.0.1 (http://rsequencing.sourceforge.net/). Data analysis was done using the RSeQC Analysis Pipeline for RNA Sequencing (MAPRSeq v3.1.3) with paired-end reads aligned using STAR-2.6.1c and alignment to the human genome reference hg38.

Fecal metabolomics. HPLC–MS. Fecal samples were deproteinized with 6x volume of cold acetonitrile/methanol (1:1 ratio), kept on ice with intermittent vortexing for 30 min at 4 °C, then centrifuged at 18,000 g, 4 °C, 15 min. Phenylalanine (3 μl at 30 μg ng⁻¹) was added as internal standard to each sample before deproteinization. The supernatants were divided into 2 aliquots and dried down for analysis on a Quadrupole Time-of-Flight mass spectrometer (Agilent Technologies, 6530 Q-TOF) coupled with an ultra high pressure liquid chromatograph (1290 Infinity II, Agilent Technologies). Fecal samples were acquired in both positive and negative electrospray ionization conditions over a mass range of 100–1,200 m/z at a resolution of 10,000–35,000 (separate runs). Metabolite separation was achieved using 2 columns of differing polarity, a hydrophilic interaction column (HILIC, ethylene-bridged hybrid 2.1 x 150 mm, 1.7 mm; Waters) and a reversed-phase C18 column (high-strength silica 2.1 x 150 mm, 1.8 mm; Waters). For each column, the run time was 20 min using a flow rate of 400 μl min⁻¹. A total of 4 runs per sample was performed to give maximum coverage of metabolites. Samples were injected in duplicate or triplicate, and a quality control sample, made up of a subset of samples from the study was injected several times during a run. All raw data files obtained were converted to compound exchange file format using MassHunter DA software and Mass Profiler Professional (Agilent) was used for data alignment and to convert each metabolite feature (m/z × intensity × time) into a matrix of detected peaks for compound identification.

Bioinformatics. An unsupervised principal component analysis, analysis of variance (ANOVA), three-dimensional plot and heat map and a partial least square discrimination analysis (PLS-DA) comparison between groups were obtained for analysis. This gives a list of accurate mass molecular weights of differentially expressed compounds that was run against the Metlin database to give putative identification (IDs). The list of components would have a putative ID or a mass (m/z) value depending on whether a match was found. Components that were assigned putative IDs were further examined by comparison to a purchased reference standard of the proposed compound. Mass accuracy of the Q-TOF method was <5 ppm, with retention time precision better than 0.2%. A 1.2-fold change was determined to be detected with a confidence level of 4%.

Raw LC–MS data files were processed using a custom in-house constructed pipeline. Data files were first segregated by the analytical mode: negative C18, positive C18, negative HILIC or positive HILIC. Each segregated file set was processed independently using the pipeline. This pipeline started by filtering out high intensity meta-features to the point that each metabolite was assigned a minimum intensity of at least 5,000 in at least 75% of the samples. Next, Metaobalyst features were utilized for data normalization, differential expression analysis and visualization. To account for intra- and inter-batch effects, metabolite data in each sample was normalized using total ion current detected in the respective sample. Normalized data were log-transformed. Transformed intensities of each metabolite observed across all samples were centred and scaled via mean centering. Processed metabolite data were analysed by a multivariate approach such as principal component analysis (PCA) to reveal data heterogeneity, groupings, outliers and trends. Hierarchical clustering analysis was performed to assess clustering of biological replicates with technical replicates and co-clustering that could differentiate with different subsets of clinical variables. Transformed intensities of each metabolite were compared between high- and low-PA groups using Student’s unpaired t-test. Resulting P values were adjusted for multiple testing using the Benjamini-Hochberg method (false discovery rate, FDR). Metabolites with FDR-adjusted P values or q-values ≤0.05 and an absolute log(fold change) ≥1.5 (value change) between two groups were considered as statistically differentially expressed and saved for further analyses.

For metabolome analysis in mice, faecal samples were collected and sent to Metabolon where global metabolic profiling was completed using the Metabolon platform.

Humanization of germ-free mice. Animal experiments were approved by the Mayo Clinic IACUC (Protocol #A0003420-18-R20). Swiss-Webster germ-free female mice were housed in flexible film isolators (CFC) and given free access to both autoclaved food and water. Mice (4–6 weeks old) were given an oral gavage of patient faecal slurry (200 μl human faecal suspension) prepared 1:1 with pre-reduced PBS in an anaerobic growth chamber, kept at 37 °C and supplied with 15% CO2, 5% H2 and 80% N2 (Coy Laboratory Products). Upon humanization, mice were housed in individually ventilated Maxiseal cages (Amri, 3–5 mice per cage) for 6 weeks to allow for microbiome engraftment. Mouse pellets were collected at baseline (pre-gavage), 6 weeks post and during various treatments. Animals were treated in accordance with the Mayo Clinic IACUC under the approved protocol #A0003420-18-R20. Swiss-Webster germ-free female mice were initially housed in flexible film isolators with free access to both autoclaved food and water. After humanization in randomized order, mice were housed in individually ventilated sterilized Maxiseal cages with autoclaved food and water, sterile woodchip bedding and an enrichment bedding stick. Mice numbers never exceeded 5 mice per cage, with an average of 3 mice per cage. Mouse received bedding and cage changes every 2 weeks, with food and water replaced as needed. Mice were kept on a 12 h on/12 h off light cycle and handled only during the daytime to avoid disruptions in circadian rhythm. Mice were kept in a temperature and humidity-controlled room with temperature ranging from 64 to 76 °F and a relative humidity between 30 and 40%. The operator was not blinded to the assigned humanization group.

Scoring of faecal pellet frequency and consistency. Humanized mice were sterilized to serve cages with a wire rack bottom. Mice were left alone for 1 h, after which the total number of pellets in each cage was counted. Data from humanized mice representative of each human microbiome were then averaged to give the number of pellets produced per hour for the respective humanized state. Pellets collected from humanized mice were scored 0–4 on the basis of a previously reported consistency scale. Scoring of faecal pellet consistency was defined as: 0, normal/formed; 1, loose but formed pellet; 2, no form, loose with some diarrhoea; 3, diarrhoea and 4, severe, watery diarrhoea. Two independent measurements from two separate researchers were recorded, from which the average pellet consistency was calculated for each mouse. Averages of each human microbiome were then reported as an overall average for each humanized state (average of averages).

Shotgun metagenomics in humanized mice. Sequencing. Upon submission, the DNA samples were quantified using a fluorimetric PicoGreen assay. The purity of the samples was also assessed via a Nanodrop. For a sample to pass quality control, it needed to quantify >0.2 ng μl⁻¹. If the samples passed quality control, they were sequenced on a TruSeq NexteraXT DNA Sample Preparation Kit. Conventional DNA (gDNA) samples were converted to Illumina sequencing libraries using Illumina’s NexteraXT DNA sample preparation kit (FC-131-1096). The NexteraXT protocol
using a 1.1 ml Z-gel micro tube (Sarstedt) by centrifugation at 10,000 rpm. Blood was collected via intra-cardiac puncture and serum was separated. Mice were removed from standard housing and fasted but were performed in R-3.6.1 (R Development Core Team). All statistical analyses have significant predictive power (R package 'Boruta' v7.0.0). All statistical analyses were performed in R-3.6.1 (R Development Core Team). Using a permutation test (999 permutations). The permutation test was based on an F-statistic of a linear model (square-root-transformed taxa normalized abundance as the response variable) as the test statistics. For mice data, differential taxa abundance analysis was performed by fitting a generalized mixed effects model (glimmPQL function in package ‘MASS’ v7.3.53) using Penalized Quasi-Likelihood (family = quasipoisson) to account for overdispersion and correlations between samples from the same donor. Log-transformed geometric mean of pair-wise ratios (GMPR) size factor was included as the offset to address sequence depth variation. To identify differentially abundant KEGG pathways associated with the variable of interest, we fit a linear mixed effects model to the arcsine square-root-transformed relative abundance data of KEGG functions (level 3). To test the association between taxa and reference PA values. Spearman correlation analysis was used. For all taxa association analysis, FDR control (Benjamini-Hochberg procedure) was used for multiple testing correction (p-adjust function in package ‘stats’ v3.6.1). FDR-adjusted P-values or (p-values < 0.10) were considered significant.

**Results.** Random forests was used for predictive modelling on the basis of the microbiota profile (genus-level relative abundance data) using default parameters of the R implementation of the algorithm (R package ‘randomForest’). Leave-one-out cross-validation was used to assess the classification accuracy in human data, and leave-one-donor-out was applied in mice data to account for potential dependency of the microbiome derived from the same donor. Receiver operating characteristic (ROC) curve (loo-curve) was used in package ‘pROC’ (v1.16.2) was used to estimate the prediction performance. The Boruta algorithm was used to select the taxa that have significant predictive power (R package ‘ Boruta’ v7.0.0). All statistical analyses were performed in R-3.6.1 (R Development Core Team).

**In vivo permeability.** Mice were removed from standard housing and fasted but allowed access to water for 2h before experimentation. Mice were then given a gavage solution (200µl) containing 3 tracer molecules: creatinine (100mg·ml⁻¹), 4kDa FITC-dextran (60 mg·ml⁻¹) and 70kDa rhodamine B-dextran (Sigma). Five hours after gavage, mice were euthanized via intracardiac injection using a 1.1 ml Z-gel micro tube (Sarstedt) by centrifugation at 10,000 rpm for 10 min at 4°C. Creatinine was measured using a COBAS c111 (Roche Diagnostics) by the Mayo Clinic Immunochromel Chemical. Serum FITC and rhodamine B-dextran were evaluated using a Synergy Mx (BioTek) plate reader at excitation of 485 and 535 nm and emission wavelengths of 528 and 627 nm, respectively.

**Tissue histology.** Small intestinal and colonic tissues were fixed in 4% paraformaldehyde, embedded in optimal cutting temperature compound, sectioned (10µm thickness) using a cryostat (Leica Biosystems) and mounted on slides. A standard hematoxylin and eosin staining protocol was followed. Slides were then graded in blinded fashion by a pathologist to assess gross differences in morphology and immune or other cellular infiltration.

**Monoclonization of germ-free mice with GUS-expressing E. coli.** E. coli containing plasmid pPS0750, a gift from Philip Poole (Addgene plasmid # 11551; http://n2t.net/addgene:11551; RRID:Addgene_11551), was grown in Luria-Bertani (LB) broth with kanamycin (20µg·ml⁻¹) and streptomycin (20µg·ml⁻¹) and incubated at 37°C overnight. Individual colonies were streaked for isolation with single colonies picked and then grown in LB broth. Overnight cultures were spun for 5 min at maximum speed and resuspended to an OD₆₀₀ of 1.0 (10⁶ cells per ml) in sterile PBS. Transformed bacteria (200µl) or DE3 competent E. coli were used to monoclonize 10-week-old female germ-free Swiss-Webster mice in a random order. Confirmation of monoclonization by either control or GUS+ E. coli was done by collecting pellets at day 7, inoculating pellets in selective media at 37°C and examining for growth overnight. Pellets from culture positive monoclonized animals were then used to assess faecal PA as described earlier.

**Unconjugated bilirubin treatment in germ-free mice.** Unconjugated bilirubin (Sigma Aldrich, B4126) was prepared (200µM) in dimethylsulfoxide (DMSO, 0.4%) immediately before use on ice in the dark. Mice were either given 200µl of unconjugated bilirubin (treated) or DMSO as a control (vehicle) once per day over the course of 3 d. After day 3 of treatment, pellets were collected and trypsin-like activity in pellets was measured.

**In vitro trypsin inhibition by unconjugated bilirubin.** Metabolites within the bilirubin pathway were tested for the ability to inhibit trypsin activity in vitro. Biliverdin (Sigma Aldrich, 30981), mesobilirubin (Frontier Speciality Chemical, MS88-9), urobilinogen (Mybiosource, MBS173206), unconjugated bilirubin (Sigma Aldrich, B4126) and conjugated bilirubin (Frontier, B850) were reconstituted to 1 mg·ml⁻¹. These stock solutions were then used to prepare a series of dilutions. Buffer (69µl; 0.046 M Tris-HCL, 0.0115 M CaCl₂, pH 8.1) was added to a black-bottom plate along with 3µl of trypsin (10ng·µl⁻¹) and 3µl of either unconjugated or conjugated bilirubin. A 100µM substrate solution of N-fluor-Tosyl-Gly-Pro-Arg 7-amido-methylcoumarin HCl was prepared in buffer as previously described and kept on ice. Substrate (75µl) was automatically dispensed into each well to give a final reaction volume of 150µl. Reactions were briefly shaken and then transferred to a Synergy Mx multi-mode microplate reader (BioTek) and read kinetically every 30 s for the course of 3 d.

**FMT.** To test the efficacy of an FMT in lowering faecal PA, at 6 weeks post humanization, high-PA humanized mice were given an oral gavage of healthy volunteer faecal slurry (200µl, 1:2 with pre-reduced PBS under anaerobic conditions). Conditions). Pellets were collected at day 3 and 1 week post FMT and assessed for faecal PA and changes in the microbiome.

**Mice-humanization with or without A. putredinis-spiked faeces.** High-PA PB-IBS patient faecal slurries were prepared as described with the following modifications. Mice were randomly divided into two groups, one receiving high-PA faecal slurry and the other given high-PA faecal slurry with A. putredinis (ATCC 29800). To generate the A. putredinis positive high-PA faecal slurry, 1 ml of an overnight culture of A. putredinis OD 1 (10⁹ cells per ml) was mixed with approximately 500 mg of high-PA faeces under anaerobic conditions. Swiss-Webster mice (4–6 weeks old) were given an oral gavage (200µl) of either high-PA faecal slurry (high-PA control) or high-PA faecal slurry ‘spiked’ with A. putredinis. Mice were housed for 6 weeks, after which pellets collected from baseline and week 6 post gavage were evaluated for changes in faecal PA. The operator was not blinded to the assigned humanization group.

**GUS inhibition in humanized mice.** The role of GUS enzymes in regulating luminal PA and systemic bioavailability was tested using two inhibitors of GUS: broad-spectrum inhibitor δ-glucurono-1,4-lactone and the loop 1 microbial specific GUS inhibitor UNC10201652. After 6 weeks post humanization of healthy humanized mice, we administered either δ-glucurono-1,4-lactone (200 mg·kg⁻¹) or UNC10201652 (9(5-morpholin-4-yl-8-piperazin-1-yl-1,2,3,4-tetrahydro-7-thia)6,9,10,11-tetraaza-benzo[c]fluorenes) obtained from Dr. Debabrata Redinho, UNC Chapel Hill (1 µmol·mg⁻¹) daily for 12 and 14 d, respectively. Control mice for both treatments were run in parallel. Pellets were collected at 3 h post treatment and on days 12 and 14 before
killing. Faecal PA was assessed as previously described and compared between the inhibitor-treated and vehicle control mice. The operator was not blinded to the assigned humanization group.

**Statistics and reproducibility.** All continuous variables are shown as means ± s.d., and as frequencies and percentages for categorical variables. Non-Gaussian distributions were assumed, and statistical tests using either two-way Mann-Whitney U test or two-way Wilcoxon signed-ranks were done to analyse data sets. When more than 2 groups were compared, non-parametric one-way ANOVA test (Kruskal-Wallis or Dunn’s multiple comparison) and two-way ANOVA for paired comparisons with post hoc comparisons (Tukey’s or Bonferroni’s) were used. All experiments were conducted in triplicate unless otherwise noted and results are provided as means ± s.d., with results considered statistically significant at P < 0.05.

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

**Data availability**

Raw data for human and mouse metagenomics are publicly available via sequence read archive (SRA) under the BioProject accession IDs PRJNA705217 and PRJNA705695, respectively. Human RNA-seq data are deposited in the Gene Expression Omnibus (GEO) under accession number GSE168759. Faecal metaproteomic data are available at PRIDE PXD025127 and descriptors are provided in Supplementary File 1. Human colonic proteomics dataset is available as Fig. 2 source data. Human and mouse metabolic raw reads are available at Metabolomics Workbench via ST002904 and ST002909, respectively (https://www.metabolomicsworkbench.org/). Source data are provided with this paper.

**Code availability**

Proteomics codes are available at https://github.com/galaxyproteomics/metaquantome, https://github.com/galaxyproteomics/tools-galaxyp/tree/master/tools/unipept and https://biocductor.org/packages/release/bioc/html/PECA. Microbiome codes are available at https://github.com/chloehu/FTIBSpaper.

Received: 29 April 2021; Accepted: 9 March 2022; Published online: 28 April 2022

**References**

1. Turk, B. Targeting proteases: successes, failures and future prospects. Nat. Rev. Drug Discov. 5, 785–799 (2006).
2. Edogawa, S. et al. Serine proteases as luminal mediators of intestinal barrier dysfunction and symptom severity in IBS. Gut 69, 62–73 (2020).
3. Cenac, N. et al. Role for protease activity in visceral pain in irritable bowel syndrome. J. Clin. Invest. 117, 636–647 (2007).
4. Denadai-Souza, A. et al. Functional proteomic profiling of secreted serine proteases in health and inflammatory bowel disease. Sci. Rep. 8, 7834 (2018).
5. Galipeau, H. J. et al. Novel fecal biomarkers that precede clinical diagnosis of ulcerative colitis. Gastroenterology https://doi.org/10.1053/j.gastro.2020.12.004 (2020).
6. Klem, F. et al. Prevalence, risk factors, and outcomes of irritable bowel syndrome after infectious enteritis: a systematic review and meta-analysis. Gastroenterology 152, 1042–1054.e1 (2017).
7. Barbara, G. et al. Rome Foundation Working Team report on post-infection irritable bowel syndrome. Gastroenterology 156, 46–58.e7 (2019).
8. Wouters, M. M. et al. Psychological comorbidity increases the risk for functional gastrointestinal symptoms. Gut 65, 1279–1288 (2016).
9. Sundin, J. L. et al. Altered faecal and mucosal microbial composition in post-infectious irritable bowel syndrome patients correlates with mucosal lymphocyte phenotypes and psychological distress. Aliment. Pharmacol. Ther. 41, 342–351 (2015).
10. Jäランス-Huovinen, J. et al. Faecal microbiota composition and host-microbe cross-talk following gastroenteritis and in postinfectious irritable bowel syndrome. Gut 63, 1737–1745 (2014).
11. Xu, D. et al. Efficacy of faecal microbiota transplantation in irritable bowel syndrome: a systematic review and meta-analysis. Am. J. Gastroenterol. 114, 1043–1050 (2019).
12. El-Salhy, M., Yatlebakk, J. G., Gilja, O. H., Brathen Kristoffersen, A. & Haussken, T. Efficacy of faecal microbiota transplantation for patients with irritable bowel syndrome in a randomised, double-blind, placebo-controlled study. Gut 69, 859–867 (2020).
13. Ford, A. C., Harris, L. A., Lacy, B. E., Quigley, E. M. M. & Moayyedi, P. Systematic review with meta-analysis: the efficacy of probiotics, prebiotics, synbiotics and antibiotics in irritable bowel syndrome. Aliment. Pharmacol. Ther. 48, 1044–1060 (2018).
14. Bohe, M., Borgstrom, A., Genell, S. & Ohlsson, K. Determination of immunoreactive trypsin, pancreatic elastase and chymotrypsin in extracts of human feces and ileostomy drainage. Digestion 27, 8–15 (1983).
15. Genell, S. & Gustafsson, B. E. Impaired enteric degradation of pancreatic endopeptidases in antibiotic-treated rats. Scand. J. Gastroenterol. 12, 801–809 (1977).
16. Vergnolle, N. Protease inhibition as new therapeutic strategy for GI diseases. Gut 65, 1215–1224 (2016).
17. Galipeau, H. J. et al. Novel role of the serine protease inhibitor elafin in gluten-related disorders. Am. J. Gastroenterol. 109, 748–756 (2014).
18. Hippel, J. P. et al. Food-grade bacteria expressing elafin protect against inflammation and restore colon homeostasis. Sci. Transl. Med. 4, 158ra144 (2012).
19. Qin, X. Inactivation of digestive proteases by deconjugated bilirubin: the possible evolutionary driving force for bilirubin or biliverdin predominance in animals. Gut 56, 1641–1642 (2007).
20. Poller, R. M. et al. An atlas of beta-glucuronidases in the human intestinal microbiome. Structure 25, 967–977.e5 (2017).
21. He, Y. et al. Bacterial beta-glucuronidase alleviates dextran sulfate sodium-induced colitis in mice: a possible crucial new diagnostic and therapeutic target for inflammatory bowel disease. Biochem. Biophys. Res. Commun. 513, 426–433 (2020).
22. Berumen, A. et al. Characteristics and risk factors of post-infection irritable bowel syndrome after Campylobacter enteritis. Clin. Gastroenterol. Hepatol. https://doi.org/10.1016/j.cgh.2020.07.033 (2020).
23. Yoon, H. et al. Increased pancreatic protease activity in response to antibiotics impairs gut barrier and triggers colitis. Cell. Mol. Gastroenterol. Hepatol. 6, 370–388.e373 (2018).
24. Blank, C. et al. Disseminating metaproteomic informatics capabilities and knowledge using the Galaxy-P framework. Proteomes https://doi.org/10.3390/proteomes6010007 (2018).
25. Easterly, C. W. et al. Metaquantome: an integrated, quantitative metaproteomics approach reveals connections between taxonomy and protein function in complex microbiomes. Mol. Cell. Proteom. 18, S82–S91 (2019).
26. Rolland-Fourcade, C. et al. Epithelial expression and function of trypsin-3 in irritable bowel syndrome. Gut 66, 1767–1778 (2017).
27. 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).
28. Flores, R. et al. Association of fecal microbial diversity and taxonomy with selected enzymatic functions. PLoS ONE 7, e39745 (2012).
29. Creekmore, B. C. et al. Mouse gut microbiome-encoded beta-glucuronidases identified using metagenome analysis guided by protein structure. mSystems https://doi.org/10.1128/mSystems.00452-19 (2019).
30. Macfadyen, A. & Ho, K. J. d-glucaric acid-1,4-lactone: its excetration in the bile and urine and effect on the biliary secretion of beta-glucuronidase after oral administration in rats. Hepatology 9, 552–556 (1989).
31. Pellock, S. J. et al. Gut microbial beta-glucuronidase inhibition via catalytic cycle interception. ACS Cent. Sci. 4, 868–879 (2018).
32. Gecce, K. et al. Increased faecal serine protease activity in diarrhoeic IBS patients: a colonic lumenal factor impairing colonic permeability and sensitivity. Gut 57, 591–599 (2008).
33. Toot, D. et al. Characterisation of faecal protease activity in irritable bowel syndrome with diarrhoea: origin and effect of gut transit. Gut 63, 753–760 (2014).
34. Carroll, J. M. & Maharshak, N. Enteric bacterial proteases in inflammatory bowel disease – pathophysiology and clinical implications. World J. Gastroenterol. 19, 7351–7453 (2013).
35. Jablouzí, A. et al. Fecal serine protease profiling in inflammatory bowel diseases. Front. Cell. Infect. Microbiol. 10, 21 (2020).
36. Caminero, A. et al. Duodenal bacterial proteolytic activity determines sensitivity to dietary antigens through protease-activated receptor-2. Nat. Commun. 10, 1198 (2019).
37. McCarville, J. L. et al. A commensal Bifidobacterium longum strain prevents gluten-related immunopathology in mice through expression of a serine protease inhibitor. Appl. Environ. Microbiol. 83, e01323-17 (2017).
38. Parker, B. J., Wearsch, P. A., Velco, A. C. M. & Rodriguez-Palacios, A. The genus Alistipes: gut bacteria with emerging implications to inflammation, cancer, and mental health. Front. Immunol. 11, 906 (2020).
Extended Data Fig. 1 | Taxa differentially abundant between low and high PA status and correlation with fecal PA.  

**a.** The microbiota of high and low PA PI-IBS patients (n = 12 high PA PI-IBS, 17 low PA PI-IBS) were compared at the phyla, class, and family levels. At both phyla and class levels of taxonomy, no significant differences were observed between the two groups of patients. However, one taxon, *Rikenellaceae*, was differential at the family level (FDR < 0.05). *Rikenellaceae*, the family of *A. putredinis* and was in higher abundance in low PA PI-IBS subjects with an average abundance of 5.3%, compared to 1.1% in high PA PI-IBS individuals.  

**b.** Correlation scatterplots for differentially abundant taxa negatively correlating with PA. In addition to the top 3, an additional 8 differentially abundant (11/14) taxa correlated with low fecal PA. They exhibited weak negative correlation (0.31 ≤ |r| ≤ 0.38). High PA (orange) and healthy volunteers (green) are plotted. Correlation coefficients (r) and q-values are generated from comparisons within the entire cohort to assess the relationship between PA and taxa abundance (n = 21 HV, 12 high PA PI-IBS, 17 low PA PI-IBS). Grey area shows 95% confidence level for linear smooths.  

**c.** Correlation scatterplots for differentially abundant taxa that positively correlate with PA. Of the 14 differential taxa, only 3 positively correlated with high fecal PA. *Clostridium clostridioforme*, *Pseudomonas* and *Ruminococcus gnavus* all had a weak, positive correlation (0.32 ≤ |r| ≤ 0.35). Plots were generated as stated above. Grey area shows 95% confidence level for linear smooths.
Extended Data Fig. 2 | See next page for caption.
Extended Data Fig. 2 | Specific serine protease activity assays in both high PA human and high PA humanized mouse fecal supernatants. a, Specific serine protease activity of human fecal samples. Using an enzyme preferential substrate assay, fecal trypsin (*p = 0.049, **p = 0.004), chymotrypsin (**p = 0.004) and pancreatic elastase (*p = 0.03, **p = 0.007) activities were increased in fecal supernatants generated from high PA individuals compared to both low PA PI-IBS and healthy volunteers (One-way ANOVA, multiple comparisons Kruskal-Wallis, n = 6 FSNs/group). b, Specific activity of serine proteases in humanized mice and germ-free mice. Mice humanized with high PA microbiota had increased chymotrypsin (*p = 0.049, **p = 0.004), pancreatic elastase (healthy **p = 0.008, low PA PI-IBS **p = 0.002) and neutrophil elastase activity (*p = 0.01, **p = 0.002) in fecal supernatants compared to either healthy volunteer or low PA PI-IBS humanized mice. Decreased trypsin (healthy **p = 0.008, low PA PI-IBS **p = 0.002), chymotrypsin (**p = 0.002), pancreatic elastase (**p = 0.002) and neutrophil elastase (*p = 0.02, **p = 0.004) activity were seen in healthy and low PA PI-IBS humanized mice compared to the germ-free (GF) state. (One-way ANOVA, multiple comparisons Kruskal-Wallis, n = 3 mice tested/humanization, 6 humanizations/phenotype, n = 6 germ-free mice). Boxplots: lower, middle and upper hinges correspond to 25th, 50th and 75th percentiles. Upper and lower whiskers extend to the largest and smallest value no further than 1.5 * IQR from the respective hinge.
Extended Data Fig. 3 | Tissue protease activity, gross morphology or histopathology of the high PA, low PA and healthy volunteer humanized mice.

a, In situ zymography for trypsin-like activity in mouse colonic tissue. No differences were observed between high PA and healthy humanized mice (n = 6 mice/group, data presented as mean ± SD).
b, Representative in situ zymography image of high PA and healthy mouse tissue. SYTOX Green Nuclear Stain (ThermoFisher, S7020) pseudocolored blue, N-p-Tosyl_Gly-Pro-Arg 7-amido-4-methylcoumarin hydrochloride cleavage pseudocolored green (Scale bar 50 µm).
c, Mouse weight, cecal weight and colonic length of humanized mice. Post-mortem weight was collected on humanized mice after which the gastrointestinal tract was removed and cecal weight was recorded. Colonic length measurements were done from proximal cecum to the distal rectum (n = 6 human feces/phenotype, dots represent average). Boxplots: lower, middle and upper hinges correspond to 25th, 50th and 75th percentiles. Upper and lower whiskers extend to the largest and smallest value no further than 1.5 * IQR from the respective hinge.
d, Histological examination of the gastrointestinal tract of humanized mice. Distal small bowel and distal colon tissue sections were evaluated by a pathologist (RG) in blinded manner. No observed differences in inflammation and presence of immune cells between humanized mice (n = 6 mice scored/phenotype, Scale bar 50 µm).
Extended Data Fig. 4 | Higher level taxonomic evaluation of human fecal samples used for humanization. Comparisons were made between the microbiota of healthy, low PA and high PA PI-IBS fecal slurries used for mouse humanization (n = 6 human feces/phenotype) which showed no significant differences at phylum, class and family levels.
Extended Data Fig. 5 | Differentially abundant taxa and KEGG pathways between healthy, low PA and high PA PI-IBS humanized mice.

**a**. Comparison between healthy and high PA PI-IBS engrafted microbiota. 32 differentially abundant taxa were identified between high PA with healthy humanized mice. Of the identified taxa, 13 were in greater fold abundance in healthy humanized mice compared to high PA PI-IBS mice. Colors denote greater abundance in the respective humanized group (green: healthy, orange: high PA PI-IBS, n = 6 feces/phenotype). Numbers labeling the taxa correspond to the labels presented in the main manuscript, Fig. 4f.

**b**. Differences in observed taxa between engrafted microbiota of low PA and high PA PI-IBS humanized mice. Microbiome analysis showed 25 differential taxa between low PA and high PA PI-IBS humanized groups, with 13 in greater abundance in low PA, and 12 in high PA. Colors denote greater abundance in the respective humanized group (blue: low PA PI-IBS, orange: high PA PI-IBS, n = 6 feces/phenotype). Numbers adjacent to taxa correspond to labels provided in main manuscript Fig. 4g.

**c**. Receiver operating curve assessing random forest ability to predict PA status based on taxa in humanized mice. The ability of random forest modelling algorithm to predict PA status based on selected taxa was assessed in humanized mice with an area under curve (AUC) of 0.914, (95% CI 0.848–0.981). Grey area shows confidence shape.

**d**. Heatmap of predicted KEGG pathway differences between high PA and healthy humanized mice (n = 6 feces/phenotype).
Extended Data Fig. 6 | See next page for caption.
Extended Data Fig. 6 | Fecal microbiome transfer in high PA humanized mice results in a compositional changes. 

**a.** Microbiome profiles of high PA humanized mice receiving either a control or an FMT with healthy microbiota (n=9 mice/group) were compared at the phyla, class, and family levels. At both phyla and class levels of taxonomy, no significant differences were observed between the two groups of mice. However, at the family level, 9 differential taxa were observed. With increased abundance of *Prevotellaceae, Eubacteriaceae, Enterobacteriaceae, Bacteroidaceae*, and *Clostridiaceae* in controls while *Odoribacteraceae, Rikenellaceae, Barnesiellaceae* and *Sutterellaceae* were more abundant in mice receiving FMT (FDR < 0.1).

**b.** Correlation scatterplots for differentially abundant taxa that negatively correlate with PA post-FMT treatment. After FMT, three bacterial species were found to negatively correlate with fecal PA, all at a q < 0.1. Taxa identified were *A. putredinis, Barnesiella intestinihominis*, and *A. obesi*. All of these taxa had strong negative correlations with PA status in mice post-FMT (0.6 ≤ |r| ≤ 0.79). Correlation coefficients (r) and q-values are generated from comparisons within FMT and control animals to assess the relationship between PA and differentially abundant taxa (n = 9 mice/group). Grey area shows 95% confidence level for linear smooths.
Extended Data Fig. 7 | *In vitro* trypsin activity is suppressed by unconjugated bilirubin, and inhibition of GUS enzymes results in increased intestinal permeability. **a.** Trypsin activity in the presence of metabolites within the bilirubin deconjugation pathway. Compared to the other metabolites used for experimentation, unconjugated bilirubin was the only metabolite that significantly inhibited trypsin activity across all concentrations examined. Data presented as ∆fluorescence/time, normalized to a trypsin-only control (2-Way ANOVA, Tukey’s multiple comparison test, n = 3 *p = 0.001, data presented as mean ± SD). **b,** Time course inhibition of trypsin activity in the presence of bilirubin metabolites (n = 3 biologically independent replicates, data presented as mean ± SD). **c,** Measurement of intestinal permeability in D-Glucaro-1,4-lactone treated humanized mice. Serum 4-kDa FITC-dextran levels were greater in healthy humanized mice treated with D-Glucaro-1,4-lactone indicating inhibition of GUS enzymes causes an increase in leak pathway permeability (2-sided Mann-Whitney, n = 4/group * p = 0.03). Boxplots: lower, middle and upper hinges correspond to 25th, 50th and 75th percentiles. Upper and lower whiskers extend to the largest and smallest value no further than 1.5 * IQR from the respective hinge. **d,** Proposed mechanism of microbial based inhibition of host proteases via the production of GUS enzymes.
Extended Data Table 1 | Demographic and clinical characteristics of study volunteers

Table 1 Demographic and clinical characteristics of study volunteers

|                         | Healthy (n=38) | High PA PI-IBS (n=19) | Low PA PI-IBS (n=33) | ANOVA |
|-------------------------|---------------|-----------------------|----------------------|-------|
| Gender (Female, Male)   | 24, 14        | 11, 8                 | 19, 14               |       |
| Age (Mean, SD)          | 43, 14        | 43, 15                | 48, 18               | 0.36  |
| BMI (Mean, SD)          | 28.16, 5.10   | 31.6, 8.45            | 29.8, 6.47           | 0.35  |
| Race (% Caucasian, total n) | 93.5, 31     | 100, 16               | 100, 29              |       |
| HADS - Anxiety (Mean, SD, n) | 2.65, 2.85, 31 | 5.38, 3.91, 16        | 5.24, 2.66, 29       | <0.001|
| HADS - Depression (Mean, SD, n) | 0.77, 1.2, 31 | 2.75, 2.98, 16        | 2.83, 2.56, 29       | <0.0001 |
| BM/Day (Mean, SD, n)    | 1.36, 0.54, 37 | 2.80, 1.16, 19        | 1.77, 0.77, 31       | <0.0001 |
| Bristol Stool Form (Mean, SD, n) | 4.07, 0.67, 37 | 5.07, 0.55, 19        | 4.27, 1.11, 31       | <0.0001 |
| Ease of Passage (Mean, SD, n) | 4.00, 0.14, 37 | 4.63, 0.56, 19        | 4.19, 0.46, 31       | <0.0001 |

*Statistical analysis: Non-parametric One-way ANOVA with Dunn’s multiple comparison post-hoc and Student’s t-test. Standard deviation (SD), Body mass index (BMI, kg/m²), Hospital anxiety and depression score (HADS; 0-7 normal, 8-10 borderline abnormal, 11-21 abnormal), Bowel movement (BM), Bristol stool form (1:constipation-7:diarrhea), Ease of passage scale (1-7), IBS subtype (C: Constipation, D: Diarrhea, M: Mixed) Symptom severity score (SSS; 75-175 mild 175-300 moderate, >300 severe), Quality of life (QOL; 34-170).
Extended Data Table 2 | Enzymatic kinetics of trypsin inhibition by metabolites in bilirubin deconjugation pathway

Table 2 Enzymatic kinetics of trypsin inhibition by metabolites in bilirubin deconjugation pathway

| Inhibitor Name | Structure | $K_{obs}$ (s$^{-1}$) | $k_2/K_1$ (M$^{-1}$s$^{-1}$) | IC50 (µM) |
|----------------|----------|---------------------|-----------------------------|----------|
| Biliverdin     | ![Structure](image1) | 0.001822            | 181.8182                    | 12.68    |
| Conjugated     | ![Structure](image2) | 0.004937            | 500                         | 4.679    |
| Bilirubin      | ![Structure](image3) | 0.006642            | 666.6667                    | 3.478    |
| Unconjugated   | ![Structure](image4) | 0.005981            | 588.2353                    | 3.862    |
| Bilirubin      | ![Structure](image5) | 0.004785            | 476.1905                    | 4.828    |

Inhibition constants were measured in a 0.046M Tris-HCl, 0.0115M CaCl$_2$, held at a pH 8.1 and 37°C. 10µg/mL of bovine pancreatic trypsin was used with the substrate N-p-Tosyl-Gly-Pro-Arg 7-amido-methylcoumarin HCl stock at 100µM. Each well has a final volume of 150µL: 3µL trypsin, 3µL of inhibitor, 69µL of buffer, 75µL of buffered substrate. The second-order inhibition rate ($k_2/K_1$ (M$^{-1}$s$^{-1}$)) was generated by carrying out a series of experiments at different concentrations of the potential inhibitory metabolite(s). Non-linear regression modelling was done to determine the IC50 values for the metabolites within the bilirubin deconjugation pathway (n=3 replicates/metabolite).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
  
  *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of any covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) and variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted. Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on [statistics for biologists](https://www.nature.com/nature/research/reports/statistics-for-biologists) contains articles on many of the points above.

Software and code

**Policy information on [availability of computer code](https://www.nature.com/nature/research/reports/software-and-code)**

**Data collection**

No software was used.

**Data analysis**

For DNA sequencing of human microbiota raw reads were processed using BBduk v38.73. Processed reads were analyzed using the SHOGUN v1.0.8 taxonomy profiler, using the BURST aligner option. Profiling of the functional content was performed using SHOGUN v1.0.8 to obtain KEGG orthology entries, modules and pathways, as well as HUMANvN2 v2.8.1 using its UniRef90 database to obtain UniRef90 gene family abundance. Functional profiling against targeted databases was also performed using HUMANvN2, using the MEGARPS peptidase database release 12.0.

RNA-seq fastq files for samples were obtained via illumina sequencing platform with QC performed on the FASTQ files using RSeQC software ([http://rseqc.sourceforge.net/](http://rseqc.sourceforge.net/)). Data analysis was done using Mayo Analysis Pipeline for RNA Sequencing (MAPRSeq v3.1.3) with paired end reads were aligned using STAR-2.6.1c and alignment to the human genome reference hg38. RSeQC version used was v3.0.1.

For mice data, differential taxon abundance analysis was analyzed by fitting a generalized mixed effects model ("glmmPQL" function in package "MASS" v7.3.53) using Penalized Quasi-Likelihood (family = "quasipoisson") to account for overdispersion and correlations between samples from the same donor. For all taxa association analysis, false discovery rate (FDR) control (Benjamini-Hochberg procedure) was used for multiple testing correction ("p.adjust" function in package "stats" v3.6.1).

Random forests was used for predictive modeling based on the microbiota profile (genus-level relative abundance data) using default parameters of the R implementation of the algorithm (R package "randomForest"). R version 3.6.1, packageVersion("boruta") 7.0.0

Image J used Java 1.8.0_172 for measurement of in situ zymography.
Graphs and statistics performed in GraphPad Prism version 9.0.1.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data
All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw data for human and mouse metagenomics are publicly available via sequence read archive (SRA) under the BioProject accession IDs PRJNA705217 and PRJNA705605s respectively.
Human RNAseq data is deposited in the Gene Expression Omnibus (GEO) under accession number GSE168759.
Fecal metaproteomic data is available at PRIDE PXD025127 and descriptors provided in Supplementary file 1.
Human colonic proteomics dataset is available as figure 2 source data.

Human and mice metabiomic raw reads are available at Metabolomics Workbench via Track ID 3093 and 3097 respectively (https://www.metabolomicsworkbench.org/). Proteomics codes can be found on https://github.com/galaxyproteomics/metacountome, https://github.com/galaxyproteomics/tools-galaxy/tree/master/tools/unipept, https://bioconductor.org/packages/release/bioc/html/PECA.html.
Microbiome codes can be found at https://github.com/chloelulu/PIBSpaper.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-list.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size  
Sample size for human data (PI-IBS n=52; healthy volunteers: 38) is based on prior study (PMID 30923071) with a goal to identify at least 18 high proteolytic activity (PA) PI-IBS patients. Within the PI-IBS cohort, high and low PA groups were determined based off 90th percentile of healthy volunteers as the cut-off.

For humanized mice experiments, 6 human fecal slurrys were used per phenotype (healthy volunteer, low PA and high PA PI-IBS) to have n=6/group for the 3 states (healthy volunteer, low PA and high PA PI-IBS). Sample sizes were chosen to ensure an accurate representation for each fecal PA group could be properly assessed. We performed an analysis of the fecal samples used for humanization and found them to compositionally and structurally different from another, but were similar when compared within the groups themselves. This is presented in the manuscript and gives confidence that the group sizes are large enough to capture inter-group differences, and be representative of the larger group as a whole.

4-9 mice were humanized for each human stool depending on the experiment conducted and were dependent on the mice available at the time of experimentation. All mice humanized with a specific fecal slurry were averaged together to provide an aggregate representation of the donor microbiome and phenotype and so that conclusions were not drawn from a single humanization in a single mouse. Humanizing multiple mice with the same human fecal microbiota and providing an aggregate analysis post humanization importantly accounts for potential microbiota engraftment differences between mice, but also avoids artificially inflated sample size by using each mice as an n. We show in the manuscript that mice humanized with any of the 3 phenotypes presented cluster with one another, and within the respective groups giving confidence that the mice are representative of the groups being tested. For comparisons between humanized groups of mice, data is presented and analyzed as an average of the averages (6 humanizations/phenotype). This approach of average of averages adds robustness to the analysis and avoids inflating sample size, making the analysis more robust.

Data exclusions  
No data was excluded from analysis.

Replication  
To ensure reproducibility of the experiments, PA measurements and protein concentration of fecal supernatants were measured in triplicate for both human and mice. Additionally, PA was assessed on several occasions to assess for biological reproducibility of fecal PA.

To account for variability and ensure reproducibility in mice, all data points provided for mouse experiments are presented as an average of averages, with the number of mice used for each experiment included in the manuscript. Group sizes were dependent on human fecal material available for humanization, however a minimum of 4-6 mice were included for each humanization. All mice were included when generating the average for a given humanization with human feces. The fecal PA of each mouse was measured individually, in triplicate, and then averaged with the mice humanized with the same human fecal slurry.
For experiments involving the use of protease inhibitors, fecal microbiome transplantation as well as the specific activity in mice, fecal PA was once again assessed in triplicate and the mice were averaged together and then reported as an averaged response to treatment. All attempts at replication were successful and consistent between experiments.

For all in vitro and in vivo experiments conducted, none of the replication attempts failed and no data was omitted from analysis. All data is available in the source data documents submitted with the manuscript.

**Randomization**

Fecal samples were collected from recruited human volunteers. Randomly selected human samples were used for fecal proteomics and metabolomics. The samples selected for humanization were randomly chosen from the larger cohorts of high PA PI-IBS, low PA PI-IBS and healthy volunteers. Randomly selected humanized states were used for transcriptomics. Similarly randomly selected humanized states were used for FMT experiments and in vivo pharmacological studies.

Healthy and PI-IBS volunteers were divided into groups based on the diagnosed disease. PI-IBS volunteers were further divided into 2 groups: high proteolytic activity (PA) and low PA based on experimental testing for PA using a FITC-casein based substrate activity assay. PI-IBS volunteers were divided into high and low PA using a 90% threshold (≥1078 BAE/mg of protein) of healthy volunteers, whereby FSNs from PI-IBS volunteers above 1078 BAE/mg of protein were defined as high PA. We used the 90% cut-off as within the healthy volunteer cohort, there is a clear bimodality in the distribution, indicating the presence of two populations. The 90% cutoff of healthy defined the right-sided distribution of the overall population as this subgroup within the healthy cohort was clearly different from the rest in their fecal PA and provided the biological and statistical rationale for categorizing PI-IBS subjects into high and low PA groups. Which group each PI-IBS individual was categorized into was random as we had, and have, no way of knowing what the fecal PA of these individuals will be prior to recruitment and testing. This definition of healthy, high and low PI-IBS was used to define all aspects of the study and define the microbiota used to humanize mice.

For experiments involving mice and humanization, mice were humanized with fecal microbiota from healthy, high or low PA PI-IBS volunteers. The investigators were not blinded to what microbiota was used for humanization. The mice were then defined based on humanization with the group the microbiota originated from. For example, high PA humanized mice received an oral gavage of microbiota from a high PA volunteer. To the best of our efforts, fecal samples used for humanization were selected at random to represent each of the groups within the previously defined healthy, low and high PA groups. Mice were randomly selected and allocated into each group, and the human fecal microbiota used for humanization, and the humanization strategy was staggered between the groups. This approach reduced the potential for any mouse breeding and batch effects.

For FMT experiments, mice with high fecal PA were used for experiments. These mice were given an oral gavage of healthy commensal microbiota to determine if FMT strategy would reduce fecal PA. The healthy, low PA microbiota used was defined as low based on the initial measurements of fecal PA from humans. To study the effect of specific microbial taxa in an FMT suppressing fecal PA, 2 different healthy microbiota were selected that had A. putredinis or lacked A. putredinis and were used for FMT experiments. For all FMT experiments, mice were randomly selected for either FMT or sham gavage. For A. putredinis microbiome spiked experiments, A. putredinis was added directly to the microbiota of a high PA fecal slurry and used for humanization. No addition of A. putredinis to the same high PA microbiota was used as a control. Mice receiving the A. putredinis spike high PA microbiota were randomly selected for each group prior to humanization.

Germ-free mice were randomly assigned to the monocolonization groups as well as unconjugated bilirubin experiments.

For GUS inhibitor experiments, mice were humanized with healthy volunteer microbiota as defined by the human cohort. Humanized mice were then randomly allocated into 4 groups: 2 groups received either D-Glucaro-1-4-Lactone, or UNC10201652 GUS inhibitor and the remaining 2 groups received a sham gavage of the vehicle used for the respective inhibitor.

**Blinding**

Investigators were blinded to the status of recruited healthy and PI-IBS volunteers as fecal PA assessments were completed post-recruitment.

Investigators were not blinded to the group allocations during humanization of germ-free mice or during treatments. Human fecal samples used for humanization were randomly selected with the determined groups based on the 90th percentile cut-off of healthy individuals. Since the goal of the study is to understand whether the microbiome plays a role in PA regulation, groupings and allocations of mice within these groups were necessary to determine changes in microbiota composition, physiology, and in vivo mechanistic studies. To the best of our efforts, fecal samples mouse allocation and treatment groups were randomly assigned and always run in parallel with controls.

**Reporting for specific materials, systems and methods**

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

| Materials & experimental systems | Methods |
|---------------------------------|---------|
| n/a                             | n/a     |
| ✗ Antibodies                    | ✗ ChIP-seq |
| ✗ Eukaryotic cell lines         | ✗ Flow cytometry |
| ✗ Palaeontology and archaeology | ✗ MRI-based neuroimaging |
| ✗ Animals and other organisms   |         |
| ✗ Human research participants   |         |
| ✗ Clinical data                 |         |
| ✗ Dual use research of concern  |         |
Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals
Swiss Webster, female mice, aged 4-6 weeks.

Animals were treated in accordance with the Mayo Clinic Institutional Animal Care and Use Committee under the approved protocol #A00003420-18-R20. Swiss Webster, germ-free female mice were initially housed in flexible film isolators with free access to both autoclaved food and water. After humanization, mice were housed in individually ventilated, sterilized Maxseal cages with autoclaved food and water, sterile woodchip bedding and an enrichment bedding of crinkled nest. Mice numbers never exceed 5 mice/cage, and on average were housed 3 mice to a cage. Mice receiving bedding and cage changes every 2 weeks, with food and water replaced as needed. Mice were kept on a 12h on/12h off light cycle and handled only during the daytime to avoid disruptions in circadian rhythm. Mice were kept in a temperature and humidity controlled room with temperature ranging from 64 to 76 °F and a relative humidity between approximately 30% and 40%.

All experiments were conducted during the day time during the 12h light cycle to ensure consistency in measurements, and controlling for confounding variables such as circadian rhythm.

Wild animals
The study did not involve wild animals.

Field-collected samples
The study did not involve samples collected from the field.

Ethics oversight
Mayo Clinic Institutional Animal Care and Use Committee approved protocol (Protocol #A00003420-18-R20)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics
A total of 52 Campylobacter jejuni PI-IBS patients defined by Rome III criteria and 38 healthy volunteers (without any GI diagnoses were recruited). Both male and female participants were recruited for the study, with participants recruited between the ages of 18 and 80 years. Individuals who have had a history of abdominal surgery (except hernia, C section, hysterectomy, appendectomy or cholecystectomy), inflammatory bowel disease, microscopic colitis, or celiac disease were excluded. Recruited volunteers were not pregnant at the time of the study. Tobacco or alcohol use during the study was prohibited. Seven days prior to study participation, these medications were prohibited: any affecting gastrointestinal transit, serotonergic agents, anti-cholinergic agents, antimuscarinics, narcotics, peppermint oil, antibiotics or new probiotics. Artificial sweeteners: SplendaTM [sucralose], Nutrasweet TM [aspartame], lactulose or manitol were also prohibited 2 days before the study started. All subjects were asked to complete the Hospital Anxiety and Depression Scale (HADS) and a 7-day bowel diary. PI-IBS patients also completed the Symptom Checklist-90 (SCL-90), IBS Symptom severity scale (IBS-SSS), IBS-quality of life (IBS-QoL) questionnaire as well as the Long Bowel Disease questionnaire (DBQ).

Recruitment
PI-IBS patients were identified through a prospective survey of individuals with history of laboratory proven C. jejuni infection and have subsequently developed PI-IBS. The IBS diagnosis is made if they have persistent symptoms for 6 months or more following the infection. Healthy volunteers were recruited through advertisements. They completed a screening questionnaire to confirm absence of any ongoing GI symptoms or prior GI diagnoses.

The Minnesota Department of Health (MDH) requires reporting of all laboratory-confirmed clinical Campylobacter cases in the state. Within 7 to 10 days of notification of Campylobacter cases, MDH conducts interviews with patients or guardians. We assessed PI-IBS development by sending surveys to patients aged 18-80 years from 6 to 9 months following a positive culture for Campylobacter species, recruiting individuals from November 1, 2011, and is ongoing. This was done by mail, or by telephone interviews for non-responders. Participants were then asked to complete both Rome III IBS questionnaire along with an IBS Symptom Severity Scale (IBS-SSS) questionnaire. Questions were asked to exclude diagnosis of other GI conditions. The identified patients were then invited for recruitment at Mayo Clinic where the study physician confirmed IBS and clarified all inclusion and exclusion criteria have been met and participants understand the instructions for collection of stool samples and metadata (bowel diary, psychological and other questionnaires).

Recruitment of patients and volunteers was completed in an unbiased manner and the categorization of these individuals into low and high PA groups was done experimentally, using a 90% cutoff as defined by the healthy volunteer group.

Ethics oversight
Mayo Clinic Institutional Review Board approved human studies and all participants provided a written informed consent (IRB protocol: 12-006579)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies
All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration
ClinicalTrials.gov identifier: NCT03266068
The presented study uses fecal samples and clinical metadata collected from the protocol listed above in ClinicalTrials.gov. The details on survey strategy that led to the diagnosis of PI-IBS are published in the epidemiological study from the group (PMID 23711004).

As a part of statewide surveillance, the Minnesota Department of Health (MDH) requires reporting of all laboratory-confirmed clinical Campylobacter cases in the state. Within 7 to 10 days of notification of Campylobacter cases, MDH conducts a telephone interview with the affected patient or their guardian to gather information on clinical symptoms of acute infection, exposures and treatment using a standard surveillance form. In order to assess PI-IBS development, we sent surveys to patients aged 18-80 years from 6 to 9 months following a positive culture for Campylobacter species. Surveys were conducted from November 1, 2011, to September 30, 2019, using postal mailings followed by telephone interviews for non-responders. Surveyed participants completed the Rome III IBS questionnaire and the IBS Symptom Severity Scale (IBS-SSS) questionnaire. Questions were asked to exclude diagnosis of other GI conditions like celiac disease, Crohn's disease, ulcerative colitis and microscopic colitis. The identified patients were then invited for recruitment at Mayo Clinic where the study physician confirmed the diagnosis of IBS, clarified that all inclusion and exclusion criteria have been met and participants understand the instructions for collection of stool samples and metadata [bowel diary, psychological and other questionnaires]. The bowel diary data was collected for the 7 days immediately preceding the collection of fecal sample. The fecal sample was frozen immediately and shipped in frozen state and stored in -80°C at Mayo Clinic until ready for use for generation of fecal supernatants for PA measurement, slurries for humanization and protein extraction for metaproteomics.

The Rome III criteria were used to define IBS and its subtypes based on the presence of abdominal pain or discomfort at least 2-3 times a month and alterations in bowel function (stool consistency or frequency) for ≥6 months to demonstrate chronicity of symptoms. Validated IBS-Symptom Severity Scale (IBS-SSS) and IBS-Quality of Life (IBS-Qol) scales were used. Established Bristol Stool Scale was used to determine stool form. For assessment of psychological symptoms, validated Hospital Anxiety and Depression scale were used. Outcomes measures of Bristol stool form, frequency, anxiety and depression scores are provided and scale/range described in Extended Data Table 1 and the references for the scales used to evaluate volunteers is included in the materials and methods.

Fecal samples were categorized as high PA and low PA based on the cut-off for healthy volunteers. 90th percentile of healthy volunteers was used as the cutoff (>1078 BAE/mg of protein).

Healthy as well as PI-IBS diagnosed volunteers were recruited for the study. The primary outcome for the study was evaluation of fecal PA and defining PI-IBS individuals as either high or low PA which came from an analysis of the distribution of PA in healthy volunteers which followed a log-normal distribution. Within the healthy volunteer cohort, there is a clear bimodality in the distribution, indicating the presence of two populations. We used a 90% threshold (>1078 BAE/mg of protein) to define the right-sided distribution of the overall population as this subgroup within the healthy cohort was clearly different from the rest in their fecal PA. As a consequence of this, the cutoff of 90% was used and provides the biological and statistical rationale for categorizing PI-IBS subjects into either high and low PA groups.

These grouping definitions were then used to define humanization in mice and the primary outcomes of fecal PA in humanized animals. Secondary outcomes in mice that were humanized include intestinal permeability, physiology, and the suppression of fecal PA by beta-glucuronidase inhibitors.