medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Spondyloarthritis, acute anterior uveitis, and Crohn's disease have both shared and distinct gut microbiota

Morgan Essex*[1,2,3], Valeria Rios Rodriguez*[3,4], Judith Rademacher[3,4,6], Fabian Proft[3,4], Ulrike Löber[1,2,3,7], Lajos Marko[1,2,3,7], Uwe Pleyer[3,5], Till Strowig[8,9,10], Jérémy Marchand[1,3,6], Jennifer A. Kirwan[2,6,11], Britta Siegmund†[,3,4], Sofia Kirke Forslund†[,1,2,3,7,12], Denis Poddubnyy†[,3,4,13]

1. Experimental and Clinical Research Center (ECRC), a cooperation of the Max-Delbrück Center and Charité–Universitätsmedizin, Berlin, Germany
2. Max-Delbrück Center for Molecular Medicine (MDC) in the Helmholtz Association, Berlin, Germany
3. Charité–Universitätsmedizin Berlin, a corporate member of Freie Universität Berlin and Humboldt Universität zu Berlin, Berlin, Germany
4. Department of Gastroenterology, Infectious Diseases and Rheumatology, Campus Benjamin Franklin, Charité–Universitätsmedizin Berlin
5. Department of Ophthalmology, Campus Virchow, Charité–Universitätsmedizin Berlin
6. Berlin Institute of Health (BIH) @ Charité, Berlin, Germany
7. German Center for Cardiovascular Research (DZHK), partner site Berlin
8. Department of Microbial Immune Regulation in the Helmholtz Center for Infection Research, Braunschweig, Germany
9. Cluster of Excellence RESIST (EXC 2155), Hannover Medical School, Hannover, Germany
10. Center for Individualized Medicine, Hannover, Germany
11. University of Nottingham School of Veterinary Medicine and Science, Loughborough, UK
12. Structural and Computational Biology Unit, EMBL, Heidelberg, Germany
13. Department of Epidemiology, German Rheumatism Research Center (DRFZ), Berlin, Germany

* These authors contributed equally to this work
† Correspondence should be addressed to sofia.forslund@mdc-berlin.de or denis.poddubnyy@charite.de

Abstract

Spondyloarthritis (SpA) is a group of immune-mediated diseases highly concomitant with non-musculoskeletal inflammatory disorders such as acute anterior uveitis and Crohn's disease. The gut microbiome represents a promising avenue to elucidate their underlying pathophysiology. Using discriminatory statistical methods to disentangle disease signals from one another and potential confounders, we characterized microbial signatures shared and distinct among these diseases for the first time, and established the baseline microbiota of a mixed prospective cohort. We identified a shared immune-mediated disease signal, represented by lower abundances of Lachnospiraceae taxa, most notably *Blautia* and *Ruminococcus gauvreauii* group. Patients with SpA were specifically enriched in *Collinsella* and *Lactobacillus* co-occurring with increased inflammation, while HLA-B27+ individuals displayed enriched *Faecalibacterium.* Our results indicate that the shared depletion of short-chain fatty acid producing Lachnospiraceae taxa may stem from different dysbiotic states previously described as enterotypes, which reflect long-term dietary and medication patterns and capture persistent, idiosyncratic variation between individuals.

Keywords

Gut microbiome, 16S rRNA, Spondyloarthritis, Crohn's Disease, Uveitis, Immune-mediated inflammatory disease

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Introduction

The human gut microbiome is a largely symbiotic, complex ecosystem of microorganisms residing on the intestinal mucosal surface. Bacterial species, mainly of the Firmicutes and Bacteroidetes phyla, have been much more widely studied than viral or fungal members and are found to confer a wide variety of metabolic and immunological functions upon the host, many of which are beneficial[1]. Despite high variation within and between individuals over time, stable satellite community compositions termed enterotypes have been observed across populations. Enterotypes represent taxonomic and functional bacterial niches that correlate with factors such as medication intake or long-term diet[2,3], and in some cases disease and systemic inflammation[4,5]. A dysbiotic microbiota composition is broadly defined as an imbalance between symbionts and pathobionts which reduces the resistance and resilience of the microbial gut ecosystem[6]. In a persistent dysbiotic state, physiological changes such as epithelial barrier integrity may become compromised and increase intestinal permeability[7]. This “leaky gut” phenomenon is thought to drive the inflammation characteristic of several immune-mediated diseases[8–10].

Spondyloarthritis (SpA) refers to one such group of immune-mediated inflammatory diseases with a complex clinical spectrum in genetically-susceptible individuals. An estimated 50-75% of all SpA patients and as many as 90% of radiographic axial SpA patients carry the human leukocyte antigen (HLA)-B27 gene[11], making the association one of the strongest ever reported between an HLA allele and a disease[12]. The clinical SpA features include inflammation of the axial skeleton and extra-musculoskeletal manifestations such as psoriasis, acute anterior uveitis (AAU), and inflammatory bowel diseases (IBD), both Crohn’s disease (CD) and ulcerative colitis (UC)[13]. Up to 45% of SpA patients present with one or more extra-musculoskeletal manifestations in the course of their disease (around 33% with AAU and up to 15% with IBD), and around 20% of IBD patients and 40% of AAU patients eventually develop SpA[14–16]. These diseases have a well-documented epidemiological association, but the underlying pathophysiological mechanisms are not yet fully understood despite decades of research. More than 25 years ago it was demonstrated that HLA-B27 transgenic rats, which spontaneously develop IBD and SpA pathologies, did not develop disease in a germ-free environment[17]. Subsequent gastrointestinal colonization with a few commensals, most notably *Bacteroides spp.*, was sufficient to trigger arthritis and colitis, suggesting a causal role of the microbiome in pathogenesis[18].

Microbial dysbiosis in IBD human cohorts has been extensively characterized in the literature. Patients with CD typically present with lower gut taxonomic diversity than healthy individuals, reliably exhibiting depletions in Clostridiales taxa, such as Ruminococcaceae and Lachnospiraceae, as well as enrichments in taxa such as *Lactobacillus*, *Enterococcus*, *Escherichia* and *Shigella*[19–22]. CD cohorts have also reported a reproducible decrease in *Faecalibacterium*, a Clostridiales genus with only one known species, *F. prausnitzii.* The controlled absence of this bacterium promoted a higher risk of postoperative ileal flare in patients with CD, thereby demonstrating a specific anti-inflammatory effect[23]. Other taxonomic shifts observed in CD patients include enrichments of specific strains of adherent-invasive bacteria, such as *Escherichia coli*, which may replicate within macrophages and promote a pro-inflammatory environment[24]. Such adherent-invasive bacteria have also been found in the gut of SpA patients displaying comparable dysbiosis to CD[9] and are associated with the alteration of the intestinal mucosal barrier[25]. In parallel,

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

recent work in experimental autoimmune uveitis mice models has demonstrated that leukocyte trafficking between gut and eye and increased gut permeability precede the ocular inflammation in AAU[26–28].

Despite the high co-occurrence of these diseases, most human microbiome studies have focused on bacterial alterations in SpA, CD, and AAU individually without exploring their concomitance. Cross-disease comparisons and meta-analyses have revealed that disease associations primarily manifest as depletions of health-associated commensals or enrichments of potentially pathogenic taxa, and that nearly half of associations observed across diverse pathologies may not be specific, but rather part of a shared, more general signal[29,30]. Furthermore, it is increasingly clear that medication regimens exert a profound impact on microbiome composition[20,31–34], and studies that fail to evaluate treatment and disease concomitance are very likely to suffer from confounding and spurious associations. In our study, we aimed to characterize a robust shared microbiota among SpA, AAU, and CD for the first time in a large human cohort, and to further resolve the relevant phenotypic and covariate associations therein.

Results

Large, concomitant disease cohorts are required to overcome heterogeneous clinical presentation of SpA, AAU, and CD

We performed 16S rRNA sequencing on the stool samples of 304 patients (79 CD, 112 AAU, and 113 axSpA patients) included in the German Spondyloarthritis Inception Cohort (GESPIC), and compared them with 63 controls without any inflammatory disorder. Demographic and clinical characteristics are detailed in Table 1. More than half of the AAU cohort and one fifth of the CD cohort presented with predominantly axial SpA, with just three cases each from the CD and AAU cohorts diagnosed as exclusively peripheral SpA. A large majority of AAU and axSpA cohort patients carried HLA-B27 while less than 10% of individuals in the CD and control groups did (Extended Data Fig. 1 and Table 1). Over 85% of the patients were naïve to biologic disease modifying anti-rheumatic drugs (bDMARDs or biologics). Regarding disease activity, CD patients had an Harvey-Bradshaw Index (HBI) of 3.3±3.8, axSpA patients had an Ankylosing Spondylitis Disease Activity Score (ASDAS) of 3.5±0.8, and 52 patients (46.4%) from the AAU cohort had active uveitis at enrollment.

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

	CD n = 79	AAU n =112	axSpA n = 113	Controls n = 63	P value
Age in years, mean ± SD	37.1 ± 12.3*	42.9 ± 13.1	37.1 ± 10.6*	38.2 ±10.4*	0.002
Male sex, n (%)	37 (46.8)$_{a,b}$	52 (46.4)$_{a}$	73 (64.6)$_{b}$	26 (41.3)$_{a}$	0.002
HLA-B27 positive, n (%)	7 (8.9)$_{a}$	87 (77.7)$_{b}$	98 (86.7)$_{b}$	5 (7.9)$_{a}$	<0.001
BMI in kg/m^2, mean ± SD	24.4 ± 4.5	25.1 ± 5.9	25.1 ± 4.3	26.2 ±5.7	0.220
Smoking status: ever smoker, n (%)	45 (57.0)$_{a}$	55 (49.1)$_{a}$	60 (53.1)$_{a}$	22 (34.9)$_{a}$	0.048
Smoking status: current smoker, n (%)	26 (32.9)$_{a,b}$	23 (20.5)$_{a,b}$	42 (37.2)$_{b}$	11 (17.5)$_{a}$	0.009
Alcohol intake, g/day, mean ± SD	3.6 ± 6.1*	2.0 ± 3.5	2.3 ± 4.5*	-	<0.001
SpA, n (%)	16 (20.3)$_{a}$	59 (52.7)c	113 (100)$_{b}$	-	<0.001
Uveitis ever, n (%)	11 (13.9)$_{a}$	112 (100)$_{b}$	26 (23.0)$_{a}$	-	<0.001
Psoriasis ever, n (%)	5 (6.3)$_{a,b}$	11 (9.8)$_{a,b}$	17 (15.0)$_{b}$	-	0.006
Inflammatory bowel diseases ever, n (%)	79 (100)$_{a}$	4 (3.6)$_{b}$	9 (8.0)$_{b}$	-	<0.001
CRP mg/l, mean ± SD	10.1 ± 25.7*	4.4 ± 7.3	13.8 ± 18.2*	1.2 ± 1.8*$^{\#\S}$	<0.001
Current NSAID treatment, n (%)	19 (24.1)$_{a}$	39 (34.8)	109 (96.5)$_{b}$	44 (69.8)$_{c}$	<0.001
Current corticosteroid treatment, n (%)	29 (36.7)$_{a}$	19 (17.0)$_{b}$	7 (6.2)$_{a}$	-	<0.001
Current csDMARD treatment, n (%)	35 (44.3)$_{a}$	6 (5.4)$_{b}$	4 (3.5)$_{b}$	-	<0.001
Current bDMARD treatment, n (%)	0	4 (3.6)	0	-	-
Naive to bDMARD treatment, n (%)	76 (96.2)$_{a}$	102 (91.1)$_{a}$	88 (77.9)$_{b}$	-	<0.001

Table 1: Clinical baseline characteristics for each cohort (CD, axSpA and AAU) and controls.
Calculated p-values represent Kruskal-Wallis test or Chi-square test as appropriate for significant difference between any two cohorts. For the cohort columns, * in superscript indicates p<0.05 vs. AAU, # indicates p<0.05 vs. axSpA, and § indicates p<0.05 vs. CD (Mann-Whitney U test). Subscript letters in cohort columns (a, b and c for CD, AAU, and axSpA, respectively) denote column proportions which do *not* differ significantly from each other at the 0.05 level (z-tests for independent proportions using Bonferroni correction).

Enterotypes capture variation in alpha and beta diversity of gut microbiota

To investigate the microbiota composition of our cohorts and phenotypes of interest, we first calculated the Shannon entropy as a measure of taxonomic (alpha) diversity and performed a multivariate principal coordinates analysis (PCoA) on genus abundances (beta diversity, see Methods). Neither cohort nor phenotype was observed as a strong factor associating with microbiota composition in either analysis, though the patient's enterotype described a large amount of between-sample variation (Fig. 1a-b). Using Dirichlet multinomial mixtures[35], our samples clustered into groups of either *Bacteroides*-dominant or *Prevotella*-dominant individuals. Though patients with CD and CD with concomitant SpA were the only groups to differ significantly from controls in their taxonomic diversity (MWU p<0.1 and p<0.05, respectively), the *Prevotella* enterotype group exhibited the lowest microbial diversity on average, yet comprised relatively few CD+ individuals (Fig. 1c-e). The *Bacteroides*-dominant individuals further separated into high- and low-*Faecalibacterium* groups (*Bacteroides* 1 and *Bacteroides* 2, respectively) in line with previous reports, despite our use of relative rather than quantitative profiling[3–5,36]. No distinct *Ruminococcus* cluster was found, though the *Bacteroides* 1 enterotype had the highest relative abundances of this taxon (Extended Data Fig. 2d).

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

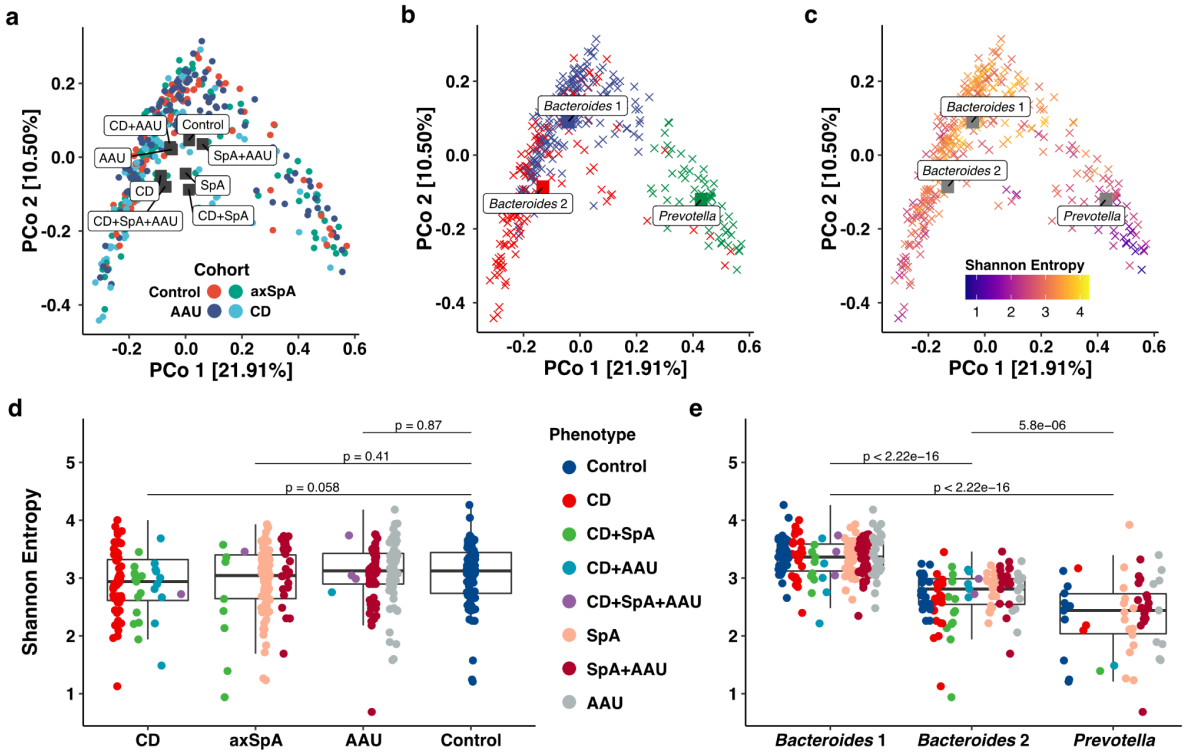

Figure 1: Enterotype effects predominate over cohort and disease phenotype effects in alpha and beta diversity analyses
a) PCoA of beta diversity calculated on pairwise Bray-Curtis dissimilarities. **b)** Dirichlet multinomial mixture modeling revealed a *Prevotella*-dominant and two *Bacteroides*-dominant clusters, where *Bacteroides* 1 is richer in *Faecalibacterium* than *Bacteroides* 2 (SFig. 1d) **c)** The *Prevotella* enterotype was characterized by low taxonomic diversity as measured by Shannon entropy. **d)** Alpha diversity was significantly lower in the CD cohort than in controls. **e)** Representation of immune-mediated phenotypes by enterotype. Taxonomic diversity was highest in the *Bacteroides* 1 enterotype, and relatively few CD+ patients belonged to the *Prevotella* enterotype. Square points in a-c indicate means of the specified groups of samples, and pairwise Mann-Whitney U tests were used in d-e.
AAU, acute anterior uveitis; CD, Crohn's disease; (ax)SpA, (axial) spondyloarthritis

Immune-mediated inflammatory diseases share a depletion of health-associated commensals

At the phylum level our disease cohorts were dominated by Firmicutes (median relative abundance 47%, 49.2% and 52.5% for CD, SpA, and AAU, respectively) followed by Bacteroidetes and Actinobacteria, except for the CD cohort which had a higher median of Proteobacteria than Actinobacteria (Extended Data Fig. 2b). Compared with controls, patients from the immune-mediated disease cohorts contained higher abundances of potential pathogen-harboring Proteobacteria and Fusobacteria phyla and lower abundances of Firmicutes and Actinobacteria on average.

To uncover taxa potentially mediating the concomitant pathology of these diseases in our cohorts, we performed a differential abundance analysis using a comprehensive modeling procedure that also allowed us to disentangle covariate effects (see Methods). We combined our cohorts into a single immune-mediated disease group (GESPIC) and compared them to the controls in order to discern taxa whose differential abundance was not specific to a single disease phenotype but rather emerged from the disease group as a whole.

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Several depletions in Firmicutes from the Lachnospiraceae family including *Blautia*, *Ruminococcus gauvreauii*, Family XII AD3011, and *Marvinbryantia,* as well as *Bifidobacterium* were robustly associated with the immune-mediated disease group as a whole rather than any single phenotype (Fig. 2a). Lachnospiraceae is an extremely diverse family of microbes highly abundant in the healthy gut microbiome[33] and among the main short-chain fatty acid (SCFA) producers in the gut[37,38]. A loss of these beneficial commensals is thus likely to co-occur with a loss of SCFA (predominantly propionate, acetate, and butyrate) known to be an energy source in the host intestine and essential for colonic health[39]. Immune-mediated patients also displayed higher abundances of *Veillonella* and *Lactobacillus,* which were positively correlated with increased inflammation as measured by C-reactive protein (CRP) and use of antibiotics within a month of sampling.

Individual disease phenotypes, inflammatory markers, and medication regimens are associated with specific gut microbial taxa

In addition to characterizing a shared disease signal, our statistical analysis allowed us to examine taxonomic abundances related to individual phenotypes and covariates of interest. As an inflammatory pathology localized in the gut itself, the CD+ phenotype exhibited the strongest dysbiosis and appeared to drive many of the shared associations, including further depletions in Ruminococcaceae and Lachnospiraceae taxa such as *Roseburia*, *Fusicatenibacter,* and the FCS020 group, all of which negatively correlated with higher inflammation. CD+ patient microbiota were also characterized by a robust enrichment of *Lachnoclostridium* and the *Clostridium innocuum* group*,* as well as *Enterococcus*, *Fusobacterium* and *Escherichia-Shigella.* With the exception of *Enterococcus,* abundances of these taxa were also naïvely associated with prior csDMARD treatment, which was most prevalent in the CD cohort (62%). Finally, higher *Fusobacterium* and *Escherichia-Shigella* abundances positively correlated with higher CRP.

We observed fewer differentially abundant taxa and smaller effect sizes in SpA+ and AAU+ patients when compared to their phenotypic absence. The SpA+ individuals exhibited a robust depletion of *Cupriavidus*, as well as an enrichment of *Collinsella* and *Lactobacillus*, likely driving the enrichment of these taxa also observed in the shared immune-mediated phenotype. *Lactobacillus* was also strongly positively correlated with higher CRP levels and the only taxon characterizing the shared signal whose enrichment appeared to be totally unrelated to enterotype-related composition (Fig. 2a). The AAU+ phenotype showed a much less pronounced depletion of Lachnospiraceae taxa and a unique enrichment of *Methanobrevibacter* and Ruminococcaceae UCG-010. An enrichment of *Coprococcus* was also observed in AAU+ patients and shared with the HLA-B27+ phenotype, except for those AAU+ patients receiving antibiotic therapy where *Coprococcus* was strongly depleted. HLA-B27+ individuals were also strongly enriched in the Lachnospiraceae NK4A136 group, *Subdoligranulum*, Ruminococcaceae UCG-009, and *Faecalibacterium*, among others (Fig. 2a).

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

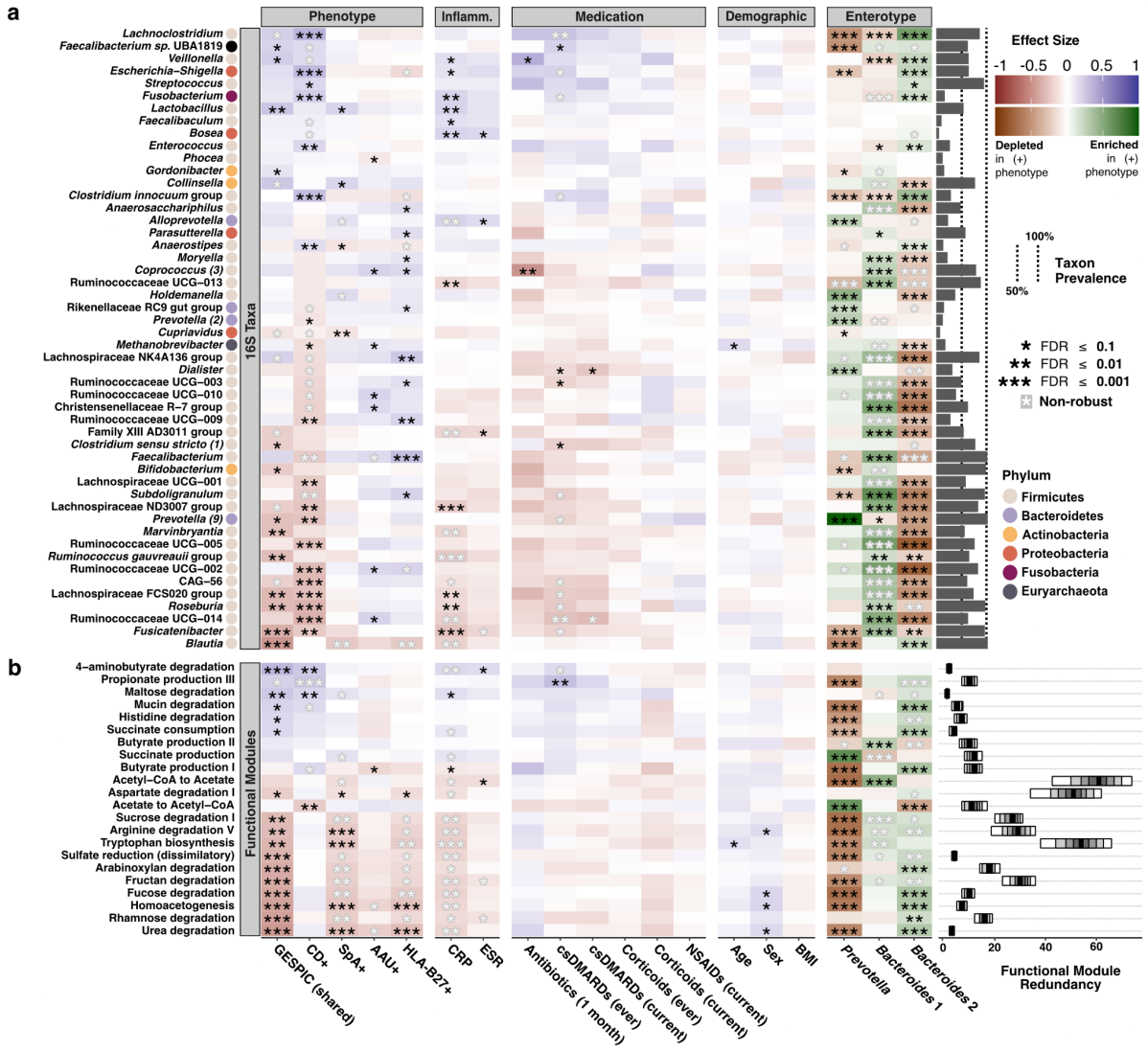

Figure 2: Specific signatures in the gut microbiota of immune-mediated disease patients reflect presence of concomitant phenotypes, inflammation, and enterotype composition

a) Key genus-level taxonomic shifts bearing robust significance to one or more phenotypes included in the leftmost panel (n=146 total genera tested after filtering, n=50 depicted). Prevalences across all 304 patients (excluding controls) shown at right. White stars denote nominal significance, i.e. from naïve association tests not considering other covariates, which our iterative nested model testing later found to explain less variation in a given taxon's abundance than another covariate in the set (see Methods). If a naïve association retained significant explanatory power across all possible models, it is shown in black and may be considered robust. Naive effect sizes (either Spearman correlations or Cliff's Delta) are depicted in the heatmap. Several taxa emerge from the shared disease cohort (GESPIC) and positively correlate with inflammation, while others co-associate with individual phenotypes and/or enterotypes. Though included as covariates in our analysis for robustness purposes, enterotypes are colored differently to reflect that their relationship to shifts in microbiome composition is fundamentally different than that of the other variables tested (i.e. they are summary representations rather than possible causal drivers, though may reflect the latter indirectly). **b)** Functional module shifts bearing robust significance to the shared phenotype or involved with SCFA and tryptophan metabolism. Module redundancy is defined as the number of genera encoding a gut metabolic module in each sample, 10-90% quantiles across all samples shown in boxplots.

AAU, acute anterior uveitis; CD, Crohn's disease; SpA, spondyloarthritis; HLA-B27, human leukocyte antigen B27; CRP, C-reactive protein in mg/L; ESR, erythrocyte sedimentation rate; csDMARDs, conventional synthetic disease-modifying antirheumatic drugs; NSAIDs, non-steroidal anti-inflammatories

Potential role of host-microbiome metabolism in immune-mediated pathologies

To allow further insight into functional metabolic alterations that might underlie taxonomic depletions observed in the immune-mediated disease group, we inferred KEGG Ortholog

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

(KO) abundances from our amplicon data and binned them into metabolic modules, which comprise key cellular enzymatic processes from manually curated reference databases (see Methods). Applying the same discriminatory differential abundance methods revealed a possible disruption in the ability of the disease group as a whole to metabolize various carbohydrates and simple sugars when compared to controls, as well as increased mucin- and urea-degrading potential. Lower microbial tryptophan biosynthesis was also predicted in our immune-mediated group, especially in SpA+ and *Prevotella*-enterotype individuals (Fig. 2b).

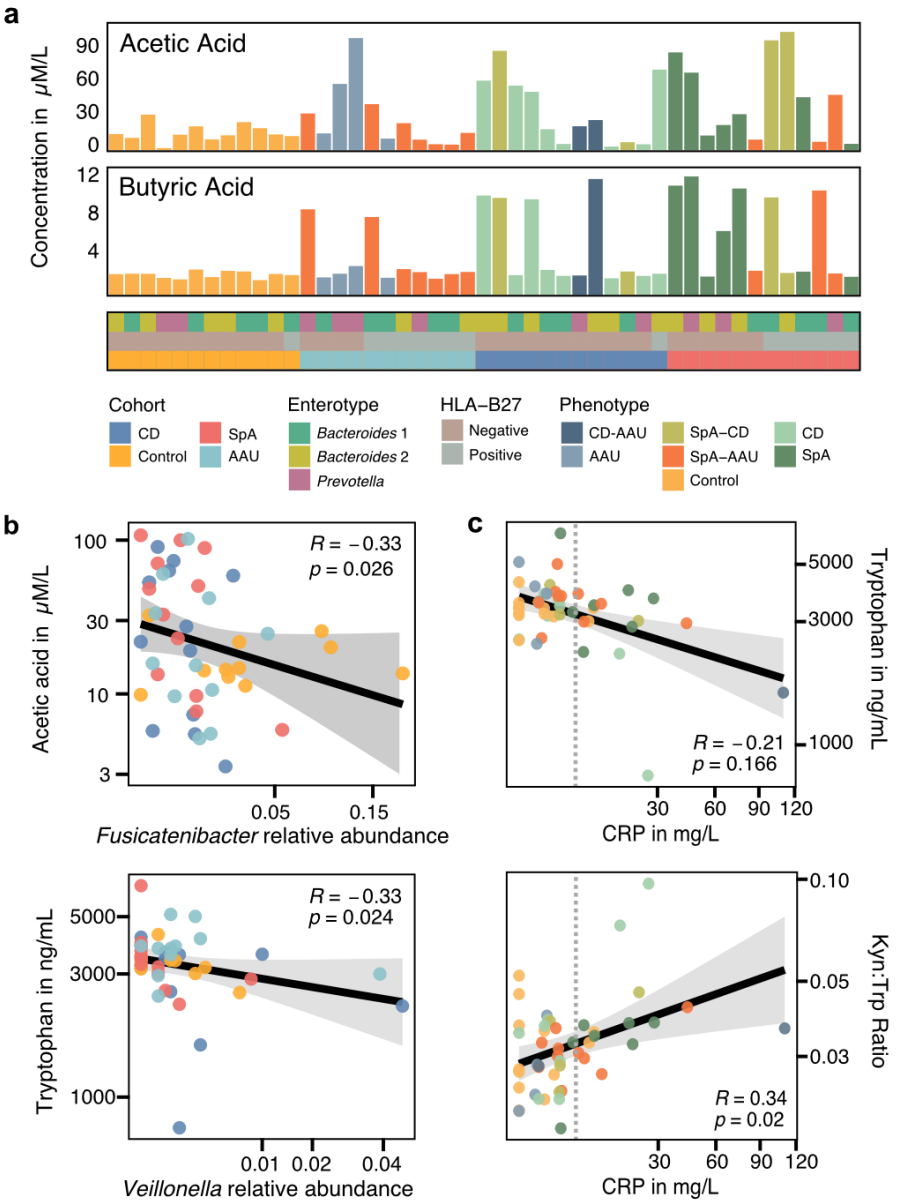

Figure 3: Targeted serum metabolomics of SCFAs and tryptophan metabolites
a) Concentrations of circulating SCFAs, which were elevated in the serum of our immune-mediated disease group compared to controls, mostly due to SpA+ individuals. **b)** Spearman correlations between relative abundances of taxa significantly related to immune-mediated disease status (low *Fusicatenibacter* and high *Veillonella*) and circulating metabolites. **c)** Spearman correlations between systemic inflammation (CRP) and immunologically-relevant metabolites or metrics, with CRP values > 5 mg/L (dashed line) implicating active disease. The kynurenine to tryptophan ratio (Kyn:Trp) positively correlated with inflammation.

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Gut inflammation and dysbiosis appear to be tightly correlated with the local balance between Tregs and pro-inflammatory Th17 cells[40], two functionally-opposing cell types that both express the aryl-hydrocarbon receptor (AhR) to integrate environmental stimuli[41], largely from microbially-produced AhR ligands, especially tryptophan metabolites[42]. To evaluate the validity of our functional predictions and better assess tryptophan metabolism, which was only sparsely profiled by our gut modules, we performed a small, targeted pilot study to quantify these metabolites as well as circulating SCFAs in twelve serum samples from each cohort. To optimize quality control in the event of continued experiments, we opted for serum rather than fecal measurements (see Methods).

Unfortunately, there was little correlation between the predicted functional module abundances and measured metabolite levels (Extended Data Fig. 3), possibly owing to imprecision in the taxonomy-to-function predictions[43], disagreement between fecal and serum metabolite levels[44,45], unmeasured confounders (e.g. diet), or low statistical power. Still, we observed higher circulating levels of butyric and acetic acids in our immune-mediated group (Fig. 3a). This was unexpected but internally consistent with our finding that higher acetic acid levels negatively correlated with relative abundances of *Fusicatenibacter,* one of the taxa most strongly depleted in these patients (Fig. 3b), and a previous study which found circulating SCFAs to be positively correlated with increased gut permeability and dysbiosis in diabetics, who had higher serum SCFAs than controls[45]. We also found *Veillonella* abundances, which were elevated in our cohorts, to be negatively correlated with circulating tryptophan, which had a weakly negative correlation with CRP. The serum kynurenine-to-tryptophan ratio, proposed as a metric of immune activation[46,47], was positively correlated with CRP (Fig. 3c), in line with a previous study which found anti-TNF-α therapy to be effective at reducing both clinical markers in SpA[48].

Discussion

CD, AAU and SpA are interrelated diseases with an established epidemiology, yet the reasons underlying their concomitance remain unclear from a pathophysiological perspective. To characterize both a shared and disease-specific gut microbiota among these immune-mediated inflammatory diseases, we assembled a large, heterogeneous cohort and performed differential abundance analysis while accounting for a large range of demographic and clinical covariates, as well as disease concomitance for the first time. Our results showed a shared depletion of predominately Lachnospiraceae taxa, most notably *Blautia,* co-occurring with increased inflammation and potential for dysregulated microbial carbohydrate metabolism in the gut. SpA+ individuals presented a robust enrichment of *Collinsella* and *Lactobacillus*, and patients with non-CD immune-mediated phenotypes displayed enriched *Faecalibacterium*.

Although it has been a subject of debate in the past[49,50], methodological advances and recent clinical literature suggest that enterotypes capture and summarize variation in microbiome composition relevant in the study of diverse chronic disease pathologies. Vandeputte *et al.* and Vieira-Silva *et al.* previously identified a *Bacteroides*-dominant, low-*Faecalibacterium* enterotype (*Bacteroides* 2) associated with inflammation (as measured by CRP and fecal calprotectin) as well as low microbial cell density in the stool[4,5]. Although we could not quantitatively profile our samples in the same manner, thus making our results less comparable and more susceptible to compositional effects, we employed the same

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

enterotyping procedure and found our own *Bacteroides* clusters to also be enriched in CD+ patients.

When we looked at each taxon independently, we found several differentially abundant taxa associated with the shared disease phenotype and several which were associated with a specific phenotype, many in conjunction with increased inflammation. The characteristic CD+ enrichments in pathogen-harboring genera such as *Fusobacterium* and *Enterococcus* were observed at a higher incidence in *Bacteroides* 2 individuals, who also displayed higher CRP and fecal calprotectin (Extended Data Fig. 4). Our data therefore sustain the hypothesis that the *Bacteroides* 2 enterotype may mediate certain dysbiotic and eventually disease states[5]; however, we also identified a *Prevotella* enterotype among our cohorts comprising individuals displaying a different dysbiosis. *Prevotella*-enterotype individuals had the lowest taxonomic diversity on average, yet they had higher abundances of taxa positively correlated with SpA+ disease activity (*Dialister,* similar to previous findings[51], and *Holdemanella*) and systemic inflammation (*Alloprevotella*). Functionally, the taxa characterizing the *Prevotella* enterotype possessed significant metagenomic potential for decreased degradation of simple sugars such as sucrose and fucose, as well as fructan (a polysaccharide), and increased potential for succinate production, which SpA+ patients also demonstrated (Fig. 2b).

Succinate is an intermediate in microbial propionate production, and its accumulation is a well-known response to hypoxia[52]. On one hand, maintenance of epithelial hypoxia has been proposed as a mechanism employed by the host immune system to influence the gut microbiota by way of favoring colonization by obligate anaerobic bacteria, which ultimately produce SCFAs; however, disrupted colonocyte metabolism causes increased epithelial oxygenation leading to succession by facultative anaerobes, including enteric pathogens[53]. On the other hand, severe epithelial hypoxia is observed in chronic gut inflammation pathologies such as CD[54]. In the same way previous studies have correlated high fat and protein diets with *Bacteroides*-dominant microbiota, and posited this enterotype as a risk factor for certain inflammatory metabolic pathologies such as CD, our findings suggest that dysregulated SCFA and carbohydrate metabolism likely to occur in *Prevotella*-dominant microbiota may be a risk factor for developing other inflammatory pathologies such as SpA, likely through similar interactions with the host immune system at the epithelial surface. Previous use of csDMARDs was robustly associated with increased potential for propionate production in our cohorts, suggesting the microbiome may mediate some of their therapeutic effects by alleviation of hypoxia-induced changes in composition, which re-establishes colonization conditions for SCFA-producing taxa. This is supported by our finding that previous csDMARD treatment was positively correlated with abundances of the obligate anaerobe *Lachnoclostridium*, whose main metabolic end product is acetate[55].

The largest taxonomic difference between our immune-mediated cohorts and controls was *Blautia*, also an anaerobic member of the Lachnospiraceae family inside the Clostridia class. It was nominally depleted in all of our phenotypes except for CD+, as well as in HLA-B27+ individuals. *Blautia* has been found in other studies to be depleted in inflammatory diseases[56], such as rheumatoid arthritis, IBD and SpA, and possesses genetic potential to produce bacteriocins and other secondary metabolites that inhibit pathogenic bacteria, such as vancomycin-resistant enterococci[57]. In a cohort of over a thousand middle-aged women, a depletion of *Blautia* was found to correlate most strongly with lower levels of indole-propionic

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

acid (IPA)[58], a byproduct of tryptophan metabolism shown to downregulate tumor necrosis factor (TNF)-α expression in enterocytes and upregulate junctional protein-coding mRNA in mice[59]. Unfortunately, we were unable to quantify IPA in our targeted metabolomics analysis. Along with other Clostridia taxa, higher abundances of *Blautia* in the gut may also have the potential to alleviate inflammation by inducing the production of immunosuppressive regulatory T cells (Tregs)[60].

We found that SpA+ patients exhibited higher abundances of *Collinsella* than controls, similar to the results obtained by Chen *et al*[61]. Other researchers showed that an enrichment of *Collinsella* reduces the expression of enterocyte tight junction proteins *in vitro*, thereby potentially contributing to gut leakage *in vivo*, and also increases the production of pro-inflammatory IL-17A and transcription factor NFkB1[62]. In healthy individuals, IL-17A, together with other subtypes of IL-17, protects the host against bacteria and fungal infection at the epithelial and mucosal barriers[63]. In SpA, an excessive activation of IL-17A contributes to chronic inflammation, as it drives an expansion of Th17 cells, which further perpetuate the production of IL-17A[64]. Similarly, colitis is associated with hyperproduction of Th17 cells, partly resulting from dysregulated NFκ-B activation responsible for regulating inflammatory T cell differentiation[65]. Our results are also consistent with lower *Faecalibacterium* abundances typically observed in CD+ cohorts[21], but the extremely strong association between this taxon and HLA-B27 suggests there is more to be uncovered about the immunomodulatory properties of certain taxa in the context of spondyloarthritis-related pathologies, especially at the molecular level. For example, previous CD-focused work isolated a 15kDA microbial anti-inflammatory molecule from *F. prausnitzii* and demonstrated its ability to inhibit NFκ-B signaling *in vitro* and to alleviate colitis in mice[66]. Our immune-mediated cohort as a whole, and especially our SpA+ individuals, displayed higher abundances of *Lactobacillus*, a purported probiotic, reflecting the complexity of host-microbiome interactions which require further characterization.

Future work which aims to elucidate the underpinnings of the microbiome in the pathophysiology of immune-mediated diseases should benefit from the present study. Using publicly-available tools and comprehensive phenotyping, our analysis allowed us to discriminate a shared microbiome signature from disease-specific associations. We performed amplicon sequencing of our cohorts, which afforded statistical power lacking in many other clinical metagenomic studies, but our insights into microbiome function were limited. Our metabolomic pilot study suggests that future work would benefit from whole genome sequencing and possibly fecal metabolite quantification to better disentangle host and microbial metabolic contributions. Although here we present only the baseline characterization of our immune-mediated cohorts, the prospective nature of our study included extensive rheumatological assessment and follow-ups in all cohorts, as well as detailed treatment and dietary information in the case of the axSpA and CD cohorts. Prospective longitudinal cohorts such as ours, paired with deep phenotyping and robust statistical methods, are crucial to identify disease-specific microbial biomarkers which are not confounded by disease concomitance or clinical factors and stay consistently enriched or depleted over time. Well-designed association studies are thus an important first step toward generating new hypotheses into host-microbiome inflammatory mechanisms, identifying microbial features which can be clinically useful as predictors for a flare, and eventually contributing to the development of new treatment or intervention targets.

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Methods

Study Design and Patient Inclusion Criteria

GESPIC is an ongoing prospective cohort initiated to study the course and long-term outcomes of SpA across its whole spectrum of clinical presentation. Since September 2015, patients have been recruited in three different arms depending on their main condition:
1) established radiographic axSpA cohort: patients were required to fulfill the modified New York criteria[67] and to be eligible to start a bDMARD therapy, by presenting high disease activity (BASDAI ≥ 4 and/or ASDAS ≥ 2.1) despite previous treatment with nonsteroidal anti-inflammatory drugs;
2) Crohn's disease cohort[68]: patients were classified according to the Montreal classification including location and behavior of CD. Gastroenterologists included recently diagnosed CD patients independent of whether they presented musculoskeletal symptoms at the time of enrollment;
3) acute anterior uveitis cohort: patients with non-infectious AAU diagnosed by an ophthalmologist were included regardless of musculoskeletal symptoms.

All patients enrolled were at least 18 years of age. Patients included in the SpA and CD cohorts were naïve to or did not receive treatment with bDMARDs for at least three months before the enrollment in the study. There were no other restrictions concerning therapy. All patients received follow-up visits every six months for the first two years at the rheumatology department at the Charité-Universitätsmedizin Berlin, afterwards yearly. They received a structured assessment of SpA manifestations, including magnetic resonance imaging of the sacroiliac joints at baseline. An experienced rheumatologist was responsible for the final diagnosis of SpA / no SpA for patients included in the CD and AAU cohorts.

Control individuals were selected from the Optiref study[69], which consisted of patients with chronic back pain who went through a standardized rheumatologic examination where the diagnosis of SpA was ruled out. Individuals with CD, AAU or psoriasis were excluded from the control group.

Clinical and biological samples were collected every six months (serum, PBMCs) alongside patient metadata (demographic information and clinical characteristics). Stool samples were collected from all patients at baseline then once per year. In this study we analyzed only the baseline timepoint. Six patients in total received antibiotics within a month of the baseline visit, eight underwent gut preparation for colonoscopy between 16-30 days before the baseline visit[70], and four patients from the AAU cohort were receiving TNF inhibitors at the baseline visit.

Sample Collection and DNA Extraction

Patients received the stool collection tube (stool collection tube with DNA Stabilizer from Stratec -PSP® Spin Stool DNA Plus Kit / PSP® Spin Stool DNA Basic Kit) on site at the baseline visit and had up to 2 weeks to collect and return the samples. Once the samples arrived at our lab, they were stored at -80°C until DNA extraction.

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

16S rRNA Sequencing

After storage at -80°C, fecal samples were defrosted on ice. Aliquots of fecal material (1ml) resuspended in RNALater were centrifuged and washed once with water to remove excess fixative and salt. Then, DNA was isolated using the ZymoBIOMICS DNA Miniprep Kit (Zymo Research) according to the manufacturer's instructions. Bead Beating was performed four times for 5 minutes each. Amplification of the V4 region (F515/R806) of the 16S rRNA gene was performed according to previously described protocols[71]. Briefly, 25 ng of DNA were used per PCR reaction (30 µl). The PCR amplification was performed using Q5 polymerase (NEB Biolabs). The PCR conditions consisted of initial denaturation for 30s at 98°C, followed by 25 cycles (10s at 98°C, 20s at 55°C, and 20s at 72°C). Each sample was amplified in triplicates and subsequently pooled. After normalization PCR amplicons were sequenced on an Illumina MiSeq platform (PE300).

Reads retrieved from 16S amplicon sequencing were analyzed using the LotuS pipeline (v1.62). The pipeline included sequence quality filtering, read merging, adapter and primer removal, chimera removal, clustering, and taxonomic classification based on the SILVA (v138), Greengenes (v13.5), and HITdb (v1.0.0) databases.

Enterotyping

Samples were enterotyped using the *DirichletMultinomial* (v1.32.0) R package according to the procedure from Holmes, Harris, and Quince[35]. The number of Dirichlet components was determined from the lowest Bayesian information criterion (BIC) on model fits ranging from 1-6 clusters. Final cluster nomenclature was based on output from the heatmapdmn function and on similarity of the resulting relative abundance profiles to those found in Vieira-Silva *et al.*[5,36] (Extended Data Fig. 2d).

Functional Annotation

Metagenome functional content was inferred from the raw amplicon sequences and processed taxonomic abundances using the PICRUSt2 method[72] (v2.3.0-b). Inferred KO counts were manually binned into KEGG Modules retaining taxonomic contributions per sample such that functional redundancy (number of genus-level metagenomes inferred to encode each module) could be calculated. To better cover SCFA-metabolism pathways of interest, KO counts were additionally binned into gut metabolic modules (GMMs)[73] using the Java implementation of the Omixer-RPM reference mapper software[74], again retaining taxonomic contribution per sample for functional redundancy calculations.

Statistical Analysis

All analyses were carried out in the R statistical programming environment (v4.0.3) using the *drake* workflow manager (v7.13.1) and *renv* package manager (v0.13.0). Raw counts were rarefied to 5000 reads and filtered to exclude OTUs not present in more than 3% of samples. The *phyloseq* package (v1.34.0) was used to bin OTUs at the genus and phylum levels (tax_glom function) and calculate the Shannon entropies, while the *vegan* (v2.5.7) and *stats* packages (vegdist and cmdscale functions) were used for beta diversity analysis. All figures were generated using *ggplot2* (v3.3.3) and *patchwork* (v1.1.1) before formatting vector files to journal specifications using Affinity Designer (v1.10).

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Statistical testing was performed using the *metadeconfoundR* package (v0.1.9) as described in Forslund *et al.*[32] and briefly described here. Cliff's delta and the Spearman correlation were used to calculate standardized effect sizes between microbial abundances and binary or continuous covariates, respectively, before FDR correction for multiple testing using the Benjamini-Hochberg method. One at a time, microbial feature abundances were rank-transformed and regressed onto a phenotypic group label (shared immune-mediated or single disease) both 1) with and 2) without a covariate, followed by a likelihood ratio test between models 1 and 2 to determine the influence of the covariate beyond that of a disease signal and vice versa, such that potentially confounded associations could be discerned. This was repeated across all available covariates and integrated to give a single robustness status (see Fig. 3). Binary phenotype labels were treated as covariates in order to disentangle their effects from the shared signal.

Metabolic Sample Selection and Statistical Analysis

We opted to quantify serum samples because fecal samples are difficult to sample consistently and reflect what is being expelled from the gut, not what the body has access to. In addition, there are standard reference materials (SRMs) that can be used for consistent quality management across batches and studies of serum. A subset of 48 serum samples were selected for tryptophan and SCFA metabolite analysis based on random subsampling within cohorts (n=12 from each of the disease groups plus controls) and using a seed for reproducibility. Relevant constraints were used for each cohort to ensure proportions of key variables observed in the larger dataset were maintained (i.e. HLA-B27 expression, disease activity, and enterotype).

Two quantitative UPLC-MS/MS analyses were carried out independently using dedicated methods that were validated following European Medicines Agency (EMA) guidelines[75]. For each, metabolite extraction was performed followed by injection on a UPLC system 1290 Infinity II (Agilent, Santa Clara, CA, USA) coupled to a TSQ Quantiva (ThermoFisher Scientific, Waltham, MA, USA)

For these analyses, commercial standards were spiked and then serially diluted into charcoal single stripped serum (Innovative Research, Novi, MI, USA) to create a calibration curve for quantification, along with solvent blanks and processed blanks. For quality management, Quality Control (QC) samples consisting of pooled samples and charcoal stripped serum samples spiked with a mix of commercial standards at known concentrations (QC standards) were also regularly injected for each batch.

The analysis of SCFA and ketone bodies (n = 11 metabolites in total) followed a previously published method[76]. In a glass vial, 40 µL of sample was incubated for 45 min at 40°C with 20 µL of 200 mM 3-Nitrophenylhydrazine hydrochloride, 20 µL of 120 mM N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride with 6 % pyridine, and 10 µL of Acetonitrile:Water (50:50) (v/v) containing stable isotope labeled internal standards (ISTDs) at 50 µM. This was followed by the addition of 410 µL of ACN:H_2O (10:90) (v/v) and centrifugation at 5,500 x *g* for 20 min at 20° C. The supernatants were then transferred to HPLC vials with inserts for UPLC-MS/MS analysis.

Liquid chromatography used a RP C18 column ACQUITY BEH 1.7 µm 2.1x100mm (Waters Corp, Milford, MA, USA) at 40 ºC for separation, with H_2O + 0.1 % (v/v) formic acid (FA) and

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

ACN + 0.1 % (v/v) FA as Phase A and Phase B, respectively. A 0.35 mL min^{-1} flow was maintained and the following gradient was used: 5 % B for 5 min, 5-55 % B in 12 min, 100 % B for 1 min, and 2 min 5 % B. Details of the Q1/Q3 transitions that were monitored can be found in Garcia-Rivera et al[76]. Injection volume was 5 µL.

The analysis of tryptophan metabolites was carried out according to an in-house developed protocol (Brüning et al, in final preparation).

Briefly, in an Eppendorf tube, 50µL of matrix was extracted with 100 µL Methanol:Water (90:10) (v/v) with 0.2% FA and 0.02% ascorbic acid containing ISTDs at 0.15 µg/mL. After one-hour incubation at -20°C to allow protein precipitation, the samples were then centrifuged at 12000 rpm for 10 minutes at 10°C. Supernatant was then transferred to HPLC vial inserts for UPLC-MS/MS analysis.

A C18 reverse phase column (Waters XSelect Premier HSS T3 Column, 100Å, 2.5 µm, 2.1 X 150 mm) was used at 15°C with the H_2O + 0.2 % (v/v) FA (Phase A) and Methanol + 0.2 % (v/v) FA (Phase B). The following elution gradient was carried out at 0.4 mL min^{-1}: 3% B for 0.45 min, 30% B at 1.2 min, 60% B at 2.7-3.75 min, 95% B at 4.5-6.6 min, 3% B at 6.75-10 min. 5µl were injected per sample.

For both methods, data were then processed with Skyline where metabolite peaks were manually identified and integrated. Area under the curve (AUC) results were exported and processed using the dedicated in-house R package MetaProc (paper in preparation). Each analyte was normalized to its respective internal standard, followed by drift correction using the notame package[77], background correction for charcoal stripped serum samples and concentration calculation based on the calibration curve.

Precision (RSD <15%) and accuracy (-/+ 15% theoretical concentration) were checked using the QC pools and QC standards. Then, for each measured sample, different statuses were automatically assigned according to calculated Limit of Detection (LOD), Lower and Upper Limits of Quantification (LLOQ, ULOQ, respectively). When calibration curve fit (R^2>98%), precision and accuracy did not match previously defined thresholds and/or when signal was lower than LOD in all samples, the analyte was not further considered.

For subsequent statistical analysis, concentration results were used only when the analyte could be quantified between LLOQ and ULOQ. When the analyte signal was detected between LOD and LLOQ for most samples, only normalized data was retained for relative comparisons. To reveal correlations with host phenotypes and microbial taxa of interest, Wilcoxon or Spearman tests were performed with valid concentration measurements while blocking for relevant covariates using the *coin* package in R (v1.4-1). Significant associations (FDR<0.05) between metabolites and the shared phenotype were independent of HLA-B27 status and each of the dichotomous phenotype labels (CD, AAU, SpA).

Ethics Declarations

All four cohorts were approved by the ethical committee (Charité-Universitätsmedizin Berlin, Berlin, Germany) and all patients gave their written informed consent. J.K. does paid consultancy for Centogene GmBH, the Rare Disease Company.

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Acknowledgements

This work was funded in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) as part of a clinical research unit (CRU339): Food allergy and tolerance (FOOD@) - Project No 428046232 and CRC-TRR241 and CRC1449 to B.S. T.S. is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2155 "RESIST" – Project ID 39087428. Additional funding was received from the German Federal Ministry for Health and Research (BMBF) as part of the WHEAT-A-BAIC consortium, and from the Berlin Institute of Health (BIH). The OptiRef[78] control cohort was partially supported from an unrestricted research grant from Novartis, and the AAU cohort by a grant from AbbVie.

Author Contributions

VRR and DP conceived the study and organized the experimental collection with help from JR, FP and UP. ME, VRR, SKF, and DP developed the hypotheses for the analysis. VRR prepared the stool samples and UL processed the sequencing data. ME and VRR, JM and JK designed the metabolomics pilot study. VRR prepared the serum metabolite samples which JM analyzed. ME conceived and performed the statistical analyses with guidance from SKF. ME and VRR interpreted the results with input from SKF, DP, and LM. ME produced the figures with input from SKF. VRR and ME wrote the manuscript with contributions from BS, SKF, and DP. All authors discussed and approved the final manuscript.

References

1 Lynch SV, Pedersen O. The Human Intestinal Microbiome in Health and Disease. N Engl J Med. 2016; **375**: 2369–2379.
2 Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA *et al.* Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes. *Science* 2011; **334**: 105–108.
3 Costea PI, Hildebrand F, Arumugam M, Bäckhed F, Blaser MJ, Bushman FD *et al.* Enterotypes in the landscape of gut microbial community composition. *Nat Microbiol* 2018; **3**: 8–16.
4 Vandeputte D, Kathagen G, D'hoe K, Vieira-Silva S, Valles-Colomer M, Sabino J *et al.* Quantitative microbiome profiling links gut community variation to microbial load. *Nature* 2017. doi:10.1038/nature24460.
5 Vieira-Silva S, Falony G, Belda E, Nielsen T, Aron-Wisnewsky J, Chakaroun R *et al.* Statin therapy is associated with lower prevalence of gut microbiota dysbiosis. *Nature* 2020; **581**: 310–315.
6 Levy M, Kolodziejczyk AA, Thaiss CA, Elinav E. Dysbiosis and the immune system. *Nat Rev Immunol* 2017; **17**: 219–232.
7 Mu Q, Kirby J, Reilly CM, Luo XM. Leaky Gut As a Danger Signal for Autoimmune Diseases. *Front Immunol* 2017; **8**: 598.
8 Manichanh C, Borruel N, Casellas F, Guarner F. The gut microbiota in IBD. Nat Rev Gastroenterol Hepatol. 2012; **9**: 599–608.
9 Breban M, Beaufrère M, Glatigny S. The microbiome in spondyloarthritis. Best Pr. Res Clin Rheumatol. 2019; **33**: 101495.
10 Ruff WE, Greiling TM, Kriegel MA. Host-microbiota interactions in immune-mediated diseases. Nat Rev Microbiol. 2020; **18**: 521–538.
11 Dashti N, Mahmoudi M, Aslani S, Jamshidi A. HLA-B*27 subtypes and their implications

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

in the pathogenesis of ankylosing spondylitis. Gene. 2018; **670**: 15–21.
12 Asquith M, Sternes PR, Costello ME, Karstens L, Diamond S, Martin TM *et al.* HLA Alleles Associated With Risk of Ankylosing Spondylitis and Rheumatoid Arthritis Influence the Gut Microbiome. Arthritis Rheumatol. 2019; **71**: 1642–1650.
13 Sieper J, Poddubnyy D. Axial spondyloarthritis. Lancet. 2017; **390**: 73–84.
14 Essers I, Ramiro S, Stolwijk C, Blaauw M, Landewé R, van der Heijde D *et al.* Characteristics associated with the presence and development of extra-articular manifestations in ankylosing spondylitis: 12-year results from OASIS. Rheumatol. Oxf. 2015; **54**: 633–40.
15 Redeker I, Siegmund B, Ghoreschi K, Pleyer U, Callhoff J, Hoffmann F *et al.* The impact of extra-musculoskeletal manifestations on disease activity, functional status, and treatment patterns in patients with axial spondyloarthritis: results from a nationwide population-based study. Ther Adv Musculoskelet Dis. 2020; **12**: 1759720x20972610.
16 Zeboulon N, Dougados M, Gossec L. Prevalence and characteristics of uveitis in the spondyloarthropathies: a systematic literature review. Ann Rheum Dis. 2008; **67**: 955–9.
17 Taurog JD, Richardson JA, Croft JT, Simmons WA, Zhou M, Fernández-Sueiro JL *et al.* The germfree state prevents development of gut and joint inflammatory disease in HLA-B27 transgenic rats. *J Exp Med* 1994; **180**: 2359–2364.
18 Rath HC, Herfarth HH, Ikeda JS, Grenther WB, Hamm TE, Balish E *et al.* Normal luminal bacteria, especially Bacteroides species, mediate chronic colitis, gastritis, and arthritis in HLA-B27/human beta2 microglobulin transgenic rats. *J Clin Invest* 1996; **98**: 945–953.
19 Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV *et al.* Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* 2012; **13**: R79.
20 Gevers D, Kugathasan S, Denson LA, Vázquez-Baeza Y, Van Treuren W, Ren B *et al.* The Treatment-Naive Microbiome in New-Onset Crohn's Disease. *Cell Host Microbe* 2014; **15**: 382–392.
21 Pascal V, Pozuelo M, Borruel N, Casellas F, Campos D, Santiago A *et al.* A microbial signature for Crohn's disease. *Gut* 2017; **66**: 813–822.
22 Franzosa EA, Sirota-Madi A, Avila-Pacheco J, Fornelos N, Haiser HJ, Reinker S *et al.* Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol* 2019; **4**: 293–305.
23 Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermudez-Humaran LG, Gratadoux J-J *et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci* 2008; **105**: 16731–16736.
24 Martinez-Medina M, Aldeguer X, Lopez-Siles M, González-Huix F, López-Oliu C, Dahbi G *et al.* Molecular diversity of Escherichia coli in the human gut: new ecological evidence supporting the role of adherent-invasive E. coli (AIEC) in Crohn's disease. Inflamm Bowel Dis. 2009; **15**: 872–82.
25 Ciccia F, Guggino G, Rizzo A, Alessandro R, Luchetti MM, Milling S *et al.* Dysbiosis and zonulin upregulation alter gut epithelial and vascular barriers in patients with ankylosing spondylitis. Ann Rheum Dis. 2017; **76**: 1123–1132.
26 Nakamura YK, Janowitz C, Metea C, Asquith M, Karstens L, Rosenbaum JT *et al.* Short chain fatty acids ameliorate immune-mediated uveitis partially by altering migration of lymphocytes from the intestine. Sci Rep. 2017; **7**: 11745.
27 Janowitz C, Nakamura YK, Metea C, Gligor A, Yu W, Karstens L *et al.* Disruption of Intestinal Homeostasis and Intestinal Microbiota During Experimental Autoimmune Uveitis. Invest Ophthalmol Vis Sci. 2019; **60**: 420–429.
28 Rosenbaum JT, Asquith M. The microbiome and HLA-B27-associated acute anterior uveitis. *Nat Rev Rheumatol* 2018; **14**: 704–713.
29 Duvallet C, Gibbons SM, Gurry T, Irizarry RA, Alm EJ. Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat Commun* 2017; **8**: 1784.
30 Wirbel J, Zych K, Essex M, Karcher N, Kartal E, Salazar G *et al.* Microbiome meta-analysis and cross-disease comparison enabled by the SIAMCAT machine

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

learning toolbox. *Genome Biol* 2021; **22**: 93.
31 Forslund K, Hildebrand F, Nielsen T, Falony G, Le Chatelier E, Sunagawa S *et al.* Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* 2015; **528**: 262–266.
32 Forslund SK, Chakaroun R, Zimmermann-Kogadeeva M, Markó L, Aron-Wisnewsky J, Nielsen T *et al.* Combinatorial, additive and dose-dependent drug–microbiome associations. *Nature* 2021; : 1–6.
33 Maier L, Pruteanu M, Kuhn M, Zeller G, Telzerow A, Anderson EE *et al.* Extensive impact of non-antibiotic drugs on human gut bacteria. Nature. 2018; **555**: 623–628.
34 Vujkovic-Cvijin I, Sklar J, Jiang L, Natarajan L, Knight R, Belkaid Y. Host variables confound gut microbiota studies of human disease. *Nature* 2020; **587**: 448–454.
35 Holmes I, Harris K, Quince C. Dirichlet Multinomial Mixtures: Generative Models for Microbial Metagenomics. *PLOS ONE* 2012; **7**: e30126.
36 Vieira-Silva S, Sabino J, Valles-Colomer M, Falony G, Kathagen G, Caenepeel C *et al.* Quantitative microbiome profiling disentangles inflammation- and bile duct obstruction-associated microbiota alterations across PSC/IBD diagnoses. *Nat Microbiol* 2019; **4**: 1826–1831.
37 Vacca M, Celano G, Calabrese FM, Portincasa P, Gobbetti M, De Angelis M. The Controversial Role of Human Gut Lachnospiraceae. *Microorganisms* 2020; **8**: 573.
38 Sorbara MT, Littmann ER, Fontana E, Moody TU, Kohout CE, Gjonbalaj M *et al.* Functional and Genomic Variation between Human-Derived Isolates of Lachnospiraceae Reveals Inter- and Intra-Species Diversity. *Cell Host Microbe* 2020; **28**: 134-146.e4.
39 Tan J, McKenzie C, Potamitis M, Thorburn AN, Mackay CR, Macia L. Chapter Three - The Role of Short-Chain Fatty Acids in Health and Disease. In: Alt FW (ed). *Advances in Immunology*. Academic Press, 2014, pp 91–119.
40 Omenetti S, Pizarro TT. The Treg/Th17 Axis: A Dynamic Balance Regulated by the Gut Microbiome. *Front Immunol* 2015; **6**: 639.
41 Li Y, Innocentin S, Withers DR, Roberts NA, Gallagher AR, Grigorieva EF *et al.* Exogenous Stimuli Maintain Intraepithelial Lymphocytes via Aryl Hydrocarbon Receptor Activation. *Cell* 2011; **147**: 629–640.
42 Gao J, Xu K, Liu H, Liu G, Bai M, Peng C *et al.* Impact of the Gut Microbiota on Intestinal Immunity Mediated by Tryptophan Metabolism. *Front Cell Infect Microbiol* 2018; **8**: 13.
43 Sun S, Jones RB, Fodor AA. Inference-based accuracy of metagenome prediction tools varies across sample types and functional categories. *Microbiome* 2020; **8**: 46.
44 Müller M, Hernández MAG, Goossens GH, Reijnders D, Holst JJ, Jocken JWE *et al.* Circulating but not faecal short-chain fatty acids are related to insulin sensitivity, lipolysis and GLP-1 concentrations in humans. *Sci Rep* 2019; **9**: 12515.
45 Zhao L, Lou H, Peng Y, Chen S, Fan L, Li X. Elevated levels of circulating short-chain fatty acids and bile acids in type 2 diabetes are linked to gut barrier disruption and disordered gut microbiota. *Diabetes Res Clin Pract* 2020; **169**: 108418.
46 Nikolaus S, Schulte B, Al-Massad N, Thieme F, Schulte DM, Bethge J *et al.* Increased Tryptophan Metabolism Is Associated With Activity of Inflammatory Bowel Diseases. *Gastroenterology* 2017; **153**: 1504-1516.e2.
47 Eryavuz Onmaz D, Sivrikaya A, Isik K, Abusoglu S, Albayrak Gezer I, Humeyra Yerlikaya F *et al.* Altered kynurenine pathway metabolism in patients with ankylosing spondylitis. *Int Immunopharmacol* 2021; **99**: 108018.
48 Eryavuz Onmaz D, Sivrikaya A, Isik K, Abusoglu S, Albayrak Gezer I, Humeyra Yerlikaya F *et al.* Altered kynurenine pathway metabolism in patients with ankylosing spondylitis. *Int Immunopharmacol* 2021; **99**: 108018.
49 Jeffery IB, Claesson MJ, O'Toole PW, Shanahan F. Categorization of the gut microbiota: enterotypes or gradients? *Nat Rev Microbiol* 2012; **10**: 591–592.
50 Gorvitovskaia A, Holmes SP, Huse SM. Interpreting Prevotella and Bacteroides as biomarkers of diet and lifestyle. *Microbiome* 2016; **4**: 15.
51 Tito RY, Cypers H, Joossens M, Varkas G, Van Praet L, Glorieus E *et al.* Brief Report: Dialister as a Microbial Marker of Disease Activity in Spondyloarthritis. Arthritis

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Rheumatol. 2017; **69**: 114–121.
52 Connors J, Dawe N, Van Limbergen J. The Role of Succinate in the Regulation of Intestinal Inflammation. *Nutrients* 2019; **11**: 25.
53 Litvak Y, Byndloss MX, Bäumler AJ. Colonocyte metabolism shapes the gut microbiota. *Science* 2018; **362**: eaat9076.
54 Konjar Š, Pavšič M, Veldhoen M. Regulation of Oxygen Homeostasis at the Intestinal Epithelial Barrier Site. *Int J Mol Sci* 2021; **22**: 9170.
55 Yutin N, Galperin MY. A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environ Microbiol* 2013; **15**: 2631–2641.
56 Salem F, Kindt N, Marchesi JR, Netter P, Lopez A, Kokten T *et al.* Gut microbiome in chronic rheumatic and inflammatory bowel diseases: Similarities and differences. *United Eur Gastroenterol J* 2019; **7**: 1008–1032.
57 Liu X, Mao B, Gu J, Wu J, Cui S, Wang G *et al.* Blautia—a new functional genus with potential probiotic properties? *Gut Microbes* 2021; **13**: 1875796.
58 Menni C, Hernandez MM, Vital M, Mohney RP, Spector TD, Valdes AM. Circulating levels of the anti-oxidant indoleproprionic acid are associated with higher gut microbiome diversity. *Gut Microbes* 2019; **10**: 688–695.
59 Venkatesh M, Mukherjee S, Wang H, Li H, Sun K, Benechet AP *et al.* Symbiotic Bacterial Metabolites Regulate Gastrointestinal Barrier Function via the Xenobiotic Sensor PXR and Toll-like Receptor 4. *Immunity* 2014; **41**: 296–310.
60 Atarashi K, Tanoue T, Oshima K, Suda W, Nagano Y, Nishikawa H *et al.* Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature* tr; **500**: 232–236.
61 Chen Z, Qi J, Wei Q, Zheng X, Wu X, Li X *et al.* Variations in gut microbial profiles in ankylosing spondylitis: disease phenotype-related dysbiosis. Ann Transl Med. 2019; **7**: 571.
62 Chen J, Wright K, Davis JM, Jeraldo P, Marietta EV, Murray J *et al.* An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis. Genome Med. 2016; **8**: 43.
63 Matsuzaki G, Umemura M. Interleukin-17 family cytokines in protective immunity against infections: role of hematopoietic cell-derived and non-hematopoietic cell-derived interleukin-17s. Microbiol Immunol. 2018; **62**: 1–13.
64 McGonagle DG, McInnes IB, Kirkham BW, Sherlock J, Moots R. The role of IL-17A in axial spondyloarthritis and psoriatic arthritis: recent advances and controversies. Ann Rheum Dis. 2019; **78**: 1167–1178.
65 Liu T, Zhang L, Joo D, Sun S-C. NF-κB signaling in inflammation. *Signal Transduct Target Ther* 2017; **2**: 1–9.
66 Quévrain E, Maubert MA, Michon C, Chain F, Marquant R, Tailhades J *et al.* Identification of an anti-inflammatory protein from *Faecalibacterium prausnitzii* , a commensal bacterium deficient in Crohn's disease. *Gut* 2016; **65**: 415–425.
67 van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. *Arthritis Rheum* 1984; **27**: 361–8.
68 Rios Rodriguez V, Sonnenberg E, Proft F, Protopopov M, Schumann M, Kredel LI *et al.* Presence of spondyloarthritis associated to higher disease activity and HLA-B27 positivity in patients with early Crohn's disease: Clinical and MRI results from a prospective inception cohort. *Joint Bone Spine* 2022; **89**: 105367.
69 Proft F, Spiller L, Redeker I, Protopopov M, Rodriguez VR, Muche B *et al.* Comparison of an online self-referral tool with a physician-based referral strategy for early recognition of patients with a high probability of axial spa. Semin Arthritis Rheum. 2020; **50**: 1015–1021.
70 Nagata N, Tohya M, Fukuda S, Suda W, Nishijima S, Takeuchi F *et al.* Effects of bowel preparation on the human gut microbiome and metabolome. *Sci Rep* 2019; **9**: 4042.
71 Gálvez EJC, Iljazovic A, Gronow A, Flavell R, Strowig T. Shaping of Intestinal Microbiota in Nlrp6- and Rag2-Deficient Mice Depends on Community Structure. *Cell Rep* 2017; **21**:

medRxiv preprint doi: https://doi.org/10.1101/2022.05.13.22275044; this version posted May 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

3914–3926.
72 Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM *et al.* PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol* 2020; **38**: 685–688.
73 Vieira-Silva S, Falony G, Darzi Y, Lima-Mendez G, Garcia Yunta R, Okuda S *et al.* Species–function relationships shape ecological properties of the human gut microbiome. *Nat Microbiol* 2016; **1**: 1–8.
74 Darzi Y, Falony G, Vieira-Silva S, Raes J. Towards biome-specific analysis of meta-omics data. *ISME J* 2016; **10**: 1025–1028.
75 van Amsterdam P, Companjen A, Brudny-Kloeppel M, Golob M, Luedtke S, Timmerman P. The European Bioanalysis Forum community's evaluation, interpretation and implementation of the European Medicines Agency guideline on Bioanalytical Method Validation. *Bioanalysis* 2013; **5**: 645–659.
76 García-Rivera MA, Fernández-Ochoa Á, Brüning U, Fritsche-Guenther R, Kirwan JA. Identification and validation of small molecule analytes in mouse plasma by liquid chromatography–tandem mass spectrometry: A case study of misidentification of a short-chain fatty acid with a ketone body. *Talanta* 2022; **242**: 123298.
77 Klåvus A, Kokla M, Noerman S, Koistinen VM, Tuomainen M, Zarei I *et al.* "Notame": Workflow for Non-Targeted LC–MS Metabolic Profiling. *Metabolites* 2020; **10**: 135.
78 Proft F, Spiller L, Redeker I, Protopopov M, Rodriguez VR, Muche B *et al.* Comparison of an online self-referral tool with a physician-based referral strategy for early recognition of patients with a high probability of axial spa. *Semin Arthritis Rheum* 2020; **50**: 1015–1021.