Shared and Distinct Gut Microbiota in Spondyloarthritis, Acute Anterior Uveitis, and Crohn’s Disease

Morgan Essex, Valeria Rios Rodriguez, Judith Rademacher, Fabian Proft, Ulrike Löber, Lajos Markó, Uwe Pleyer, Till Strowig, Jérémy Marchand, Jennifer A. Kirwan, Britta Siegmund, and Denis Poddubnyy

Objective. Spondyloarthritis (SpA) is a group of immune-mediated diseases highly comonitant with nonmusculoskeletal inflammatory disorders, such as acute anterior uveitis (AAU) and Crohn’s disease (CD). The gut microbiome represents a promising avenue to elucidate shared and distinct underlying pathophysiology.

Methods. We performed 16S ribosomal RNA sequencing on stool samples of 277 patients (72 CD, 103 AAU, and 102 SpA) included in the German Spondyloarthritis Inception Cohort and 62 back pain controls without any inflammatory disorder. Discriminatory statistical methods were used to disentangle microbial disease signals from one another and a wide range of potential confounders. Patients were naive to or had not received treatment with biological disease-modifying antirheumatic drugs (DMARDs) for >3 months before enrollment, providing a better approximation of a true baseline disease signal.

Results. We identified a shared, immune-mediated disease signal represented by low abundances of Lachnospiraceae taxa relative to controls, most notably Fusicatenibacter, which was most abundant in controls receiving nonsteroidal antiinflammatory drug monotherapy and implied to partially mediate higher serum C-reactive protein. Patients with SpA showed an enrichment of Collinsella, whereas human leukocyte antigen (HLA)-B27+ individuals displayed enriched Faecalibacterium. CD patients had higher abundances of a Ruminococcus taxon, and previous conventional/synthetic DMARD therapy was associated with increased Akkermansia.

Conclusion. Our work supports the existence of a common gut dysbiosis in SpA and related inflammatory pathologies. We reveal shared and disease-specific microbial associations and suggest potential mediators of disease activity. Validation studies are needed to clarify the role of Fusicatenibacter in gut-joint inflammation, and metagenomic resolution is needed to understand the relationship between Faecalibacterium commensals and HLA-B27.

INTRODUCTION

The human gut microbiome is a largely symbiotic, complex ecosystem of microorganisms residing on the intestinal mucosal surface. Bacterial microbiome members, mainly of the Firmicutes and Bacteroidetes phyla, have been more widely studied than viral or fungal members and confer many beneficial metabolic and immunological functions upon the host (1). A dysbiotic microbiota composition is broadly defined as an imbalance between symbionts and pathobionts that reduces the resistance and

Supported by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) as part of a clinical research unit (CRU339); Food allergy and tolerance (FOOD8–Project ID 428046232) and CRC1449 (to Dr. Forstlund), as well as CRC-TRR241 and CRC1449 (to Dr. Sigmund). Additional funding was received from the German Federal Ministry for Health and Research (BMF) as part of the WHEAT-ABAIC consortium and from the Berlin Institute of Health (BIH). The OptRef control cohort was partially supported by an unrestricted research grant from Novartis and the acute anterior uveitis cohort by an unrestricted grant from AbbVie. Dr. Rademacher is a participant in the BIH-Charité Clinician Scientist Program funded by Charité–Universitätsmedizin Berlin and the BIH. Dr. Strowig’s work was supported by the DFG under Germany’s Excellence Strategy–EXC 2155 “RESIST” (Project ID 390874280).

1Morgan Essex, MSc: Experimental and Clinical Research Center (ECRC); a cooperation of the Max Delbrück Center and Charité–Universitätsmedizin, Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), and Charité–Universitätsmedizin Berlin (a corporate member of Freie Universität Berlin and Humboldt–Universität zu Berlin), Berlin, Germany; Valeria Rios Rodriguez, MD, Fabian Proft, MD, Britta Siegmund, MD: Medical Department of Gastroenterology, Infectiology and Rheumatology, Campus Benjamin Franklin, Charité–Universitätsmedizin Berlin, Berlin, Germany; Judith Rademacher, MD: Medical Department of Gastroenterology, Infectiology and Rheumatology, Campus Benjamin Franklin, Charité–Universitätsmedizin Berlin, Berlin, Germany; Ulrike Löber, MD, Lajos Markó, MD: ECRC, MDC, Charité–Universitätsmedizin Berlin and German Center for Cardiovascular Research (DZHK), Berlin, Germany; Uwe Pleyer, MD: Department of Ophthalmology, Campus Virchow, Charité–Universitätsmedizin Berlin, Berlin, Germany; Till Strowig, PhD: Helmholtz Center for Infection Research, Braunschweig, Germany, and Cluster of Excellence RESIST (EXC 2155), Hannover Medical School and Center for Individualized Infection Medicine (CiIM; a joint venture between the Helmholtz Center for Infection Research and the Hannover Medical School), Hannover, Germany; Jérémy Marchand, PhD: MDC and BIH Metabolomics Platform at Charité–Universitätsmedizin Berlin, Berlin, Germany; Jennifer A. Kirwan, PhD:
resilience of the microbial gut ecosystem (2). In a persistent dysbiotic state, physiological conditions such as epithelial barrier integrity may become compromised and increase intestinal permeability. This “leaky gut” phenomenon is thought to drive the inflammation characteristic of several immune-mediated diseases (3).

Spondyloarthritis (SpA) refers to one such group of immune-mediated inflammatory diseases with a complex clinical spectrum in genetically susceptible individuals. An estimated 50% to 75% of all patients with SpA and as many as 90% of patients with radiographic axial SpA carry the human leukocyte antigen (HLA)-B27 gene, making the association one of the strongest ever reported between an HLA allele and a disease (4). The clinical SpA features include inflammation of the axial skeleton and extramuscloskeletal manifestations such as psoriasis, acute anterior uveitis (AAU), and inflammatory bowel diseases (IBDs), both Crohn’s disease (CD) and ulcerative colitis (UC) (5). Up to 45% of patients with SpA present with one or more extramuscloskeletal manifestations in the course of their disease (~33% with AAU and up to 15% with IBD), and ~20% of patients with IBD and 40% of patients with AAU eventually develop SpA (6–8).

These diseases have a well-documented epidemiological association, but the underlying pathophysiological mechanisms are not yet fully understood despite decades of research. It was postulated >30 years ago that reactive arthritis (in the SpA disease family and sharing the genetic link to HLA-B27) could be triggered by antecedent gastrointestinal infection, after bacterial lipopolysaccharides were isolated from synovial fluid (9). Also, >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 (10). Subsequent gastrointestinal colonization with a few commensals was sufficient to trigger arthritis and colitis, suggesting a causal role of the microbiome in (shared) pathogenesis (11).

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 nearly half of microbial associations observed across diverse pathologies may not be specific but rather part of a shared, more general disease signal (12,13). Furthermore, it is increasingly clear that medication regimens exert a profound impact on microbiome composition and that studies that fail to account for treatment and disease concomitance are very likely to suffer from confounding and spurious associations with disease states (14–18). 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 relevant phenotypic and covariate associations therein.

**METHODS**

Patient and public involvement and patient inclusion criteria. Neither patients nor the public were involved in the design, conduct, reporting, or dissemination plans of our research. The German Spondyloarthritis Inception Cohort (GESPIQ) is an ongoing prospective cohort initiated to study the course and long-term outcomes of SpA across its whole spectrum of clinical presentation, including but not limited to gut microbiome composition. In this study, we cross-sectionally analyzed only the baseline data. Serum, stool, and peripheral blood mononuclear cell samples, demographic information, and clinical characteristics were collected.

Since September, 2015, patients have been recruited in three different arms depending on their main condition: 1) established radiographic axial SpA—patients were required to fulfill the modified New York criteria and be eligible to start biological disease-modifying antirheumatic drug (bDMARD) therapy by presenting with high disease activity [Bath Ankylosing Spondylitis Activity Index (BASDAI) ≥4 and/or Ankylosing Spondylitis Disease Activity Score (ASDAS) ≥2.1] despite previous treatment with nonsteroidal antiinflammatory drugs (NSAIDs); 2) CD (19)—patients were classified according to the Montreal classification, including location and behavior of CD, and had been recently diagnosed; 3) AAU (20)—patients with noninfectious AAU diagnosed by an ophthalmologist. CD and AAU patients were enrolled regardless of musculoskeletal symptoms, and an experienced rheumatologist was responsible for the final diagnosis of SpA or no SpA for patients included in these cohorts.

All cohorts were approved by the ethical committee (Charité–Universitätsmedizin Berlin, Berlin, Germany). All patients enrolled were at least 18 years of age and gave their written informed consent. Patients included in this analysis were naive to or had not received treatment with bDMARDs for at least 3 months before the enrollment in the study, nor had they received systemic antibiotics for at least 1 month before their baseline stool sample. There were no other restrictions concerning therapy.

Control individuals were selected from the OptiRef study, which consisted of patients with chronic back pain who went through a standardized rheumatologic examination in which the

---

*Authors Essex and Rodriguez contributed equally to this work.


Address correspondence via email to Sofia K. Forslund, PhD, at sofa.forslund@mdc-berlin.de; or to Dennis Poddubnyy, MSc, MD, at denis.poddubnyy@charite.de.

Submitted for publication September 18, 2022; accepted in revised form July 13, 2023.
diagnosis of SpA was ruled out (21). Individuals with CD, AAU, or psoriasis were excluded from the control group.

16S rRNA amplicon sequencing. After storage at −80°C, fecal samples were defrosted on ice. Aliquots of fecal material (1 ml) resuspended in RNAlater were centrifuged and washed once with water to remove excess fixative and salt. Then, DNA was isolated using the ZymoBIOICS DNA Miniprep Kit (Zymo Research) according to the manufacturer’s instructions. Bead Beating was performed 4 times for 5 minutes each. Amplification of the V4 region (F515/R806) of the 16S ribosomal RNA (rRNA) gene was performed according to previously described protocols across two separate experiments (22). Briefly, 25 ng of DNA were used per polymerase chain reaction (PCR; 30 μl). The PCR amplification was performed using Q5 polymerase (NEB Biolabs). The PCR conditions consisted of initial denaturation for 30 s at 98°C, followed by 25 cycles (10 s at 98°C, 20 s at 55°C, and 20 s at 72°C). Each sample was amplified in triplicates and subsequently pooled. After normalization, PCR amplicons were sequenced on an Illumina MiSeq platform (PE300).

Taxonomic profiling and preprocessing. Computational analysis was carried out in R (v4.0.3). Demultiplexed reads were analyzed using the DADA2 pipeline (v1.25.2), and taxonomic assignment was based on the SILVA rRNA database (v138.1) at 80% bootstrapped confidence. The DECIPHER package was used for multiple sequence alignment, and the phangorn package was used to build a phylogenetic tree. Samples failing to meet inclusion criteria, lacking important clinical metadata, or not passing quality control (ie, total read counts <5,000 and >38,000) were excluded from further analysis (Supplementary Figure 1A, available on the Arthritis & Rheumatology website at https://onlinelibrary.wiley.com/doi/10.1002/art.42658), resulting in 339 samples. Libraries were rarefied to 5,000 reads to account for variability in sequencing depth, resulting in 2,066 amplicon sequence variants (ASVs) present in at least one sample. The phylloseq::tax_glom function was implemented in a custom manner to bin ASVs.

Alpha and beta diversity analysis and enterotyping. The phylloseq::estimate_richness function was applied to raw sample counts to calculate the Shannon entropy as a measure of alpha diversity (within-sample variation). To assess beta diversity (between-sample taxonomic variation), the vegan::vegdist function was applied to rarefied ASV profiles (after total-sum scaling normalization) to calculate pairwise Bray-Curtis dissimilarity matrices, and the stats::cmdscale function was used to perform a principal coordinates ordination analysis (PCoA). The DirichletMultinomial R package was used to enterotype samples from rarefied and binned counts according to the procedure from Holmes, Harris, and Quince (23).

Differential abundance analysis. ASVs not present in more than 5% of samples or displaying mean relative abundances $<1 \times 10^{-5}$ across all samples were filtered out, resulting in $n = 442$ ASVs and $n = 123$ binned higher-level taxa (annotated at either genus- or family-level; see Supplementary Table 1, available on the Arthritis & Rheumatology website at https://onlinelibrary.wiley.com/doi/10.1002/art.42658). To examine a potential shared disease signature, all $n = 277$ disease samples (“GES-PIC”) were pooled and compared with the controls in an integrated differential abundance and association testing procedure. Samples were also grouped according to phenotype, and phenotypes with $n > 40$ samples per group (ie, SpA only, CD only, AAU only, and SpA + AAU) were subjected to this procedure to resolve phenotype-specific associations. Please see the Supplementary Methods (available at https://onlinelibrary.wiley.com/doi/10.1002/art.42658) for implementation details.

Mediation analysis. Keeping the same sets of taxa and phenotypic groups described above, nonparametric effect sizes (Cliff’s delta and the Spearman’s correlation) were calculated for each possible pairing of taxonomic and clinical features (see Supplementary Table 2 for a complete list, available on the Arthritis & Rheumatology website at https://onlinelibrary.wiley.com/doi/10.1002/art.42658), then tested for significance independent of a disease signal using blocked Wilcoxon or Spearman tests from the coin package (formula relative_abundance $\sim$ covariate | disease_status), and adjusted for multiple testing (Benjamini-Hochberg procedure).

The mdt_simple() function from the JSMediation package was used as described in Yzerbyt et al (24) to test specific causal mediation hypotheses. This approach used linear regression to estimate the paths (a, b, c, and c’) along a possible causal pathway involving three variables. The joint significance of these paths was assessed to determine whether mediation was present and, if so, what the direct and indirect contributions within the pathway were. To lower the burden of multiple testing, only configurations with significantly disease-associated taxa and covariates emerging from the above testing were considered. Please see the Supplementary Methods (available on the Arthritis & Rheumatology website at https://onlinelibrary.wiley.com/doi/10.1002/art.42658) for details on the confounder analysis with medication.

Processed data tables mentioned in the supplemental file are hosted at https://github.com/sxmorgan/gespic-public. All software and R package versions are denoted in the renv.lock file also available in the repository.

RESULTS

Clinical presentation of SpA, AAU, and CD cohorts. Demographic and clinical characteristics of our 277 patients (72 with CD, 103 with AAU, and 102 with axial SpA) and 62 back pain controls are detailed in Table 1. More than half of the AAU
the genus level, we observed substantially more variation onlinelibrary.wiley.com/doi/10.1002/art.42658). In patients in the SpA cohort clearly showed high systemic inflammation. Only 10% of individuals in the CD cohort presented with predominantly axial SpA, whereas less than 10% of individuals in the AAU and control groups did (Table 1; Supplementary Figure 1B, Table 1. Clinical baseline characteristics for each cohort—CD, axial SpA, and AAU—and controls*

<table>
<thead>
<tr>
<th></th>
<th>CD (n = 72)</th>
<th>AUU (n = 103)</th>
<th>Axial SpA (n = 102)</th>
<th>Controls (n = 62)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, mean ± SD</td>
<td>37.4 ± 12.7</td>
<td>42.8 ± 13.1</td>
<td>36.9 ± 10.6</td>
<td>38.4 ± 10.4</td>
<td>0.003†</td>
</tr>
<tr>
<td>Male sex, n (%)</td>
<td>34 (47.2)</td>
<td>48 (46.6)</td>
<td>65 (63.7)</td>
<td>26 (41.9)</td>
<td>0.015</td>
</tr>
<tr>
<td>HLA-B27 positive, n (%)</td>
<td>6 (8.3)</td>
<td>81 (78.6)</td>
<td>89 (87.3)</td>
<td>5 (8.1)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BMI in kg/m², mean ± SD</td>
<td>24.4 ± 4.6</td>
<td>24.7 ± 5.7</td>
<td>25.1 ± 4.2</td>
<td>25.9 ± 5.0</td>
<td>0.153†</td>
</tr>
<tr>
<td>Smoking status: current smoker, n (%)</td>
<td>24 (33.3)</td>
<td>22 (21.4)</td>
<td>38 (37.3)</td>
<td>11 (17.7)</td>
<td>0.02</td>
</tr>
<tr>
<td>Alcohol intake, g/day, mean ± SD</td>
<td>3.6 ± 6.3</td>
<td>2.1 ± 3.6</td>
<td>2.1 ± 4.2</td>
<td>2.1 ± 4.2</td>
<td>0.14†</td>
</tr>
<tr>
<td>SpA, n (%)</td>
<td>12 (16.7)</td>
<td>55 (53.4)</td>
<td>102 (100)</td>
<td>0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Uveitis ever, n (%)</td>
<td>11 (15.3)</td>
<td>103 (100)</td>
<td>22 (21.6)</td>
<td>0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Psoriasis ever, n (%)</td>
<td>4 (5.6)</td>
<td>10 (9.7)</td>
<td>17 (16.7)</td>
<td>0</td>
<td>0.002</td>
</tr>
<tr>
<td>CD ever, n (%)</td>
<td>72 (100)</td>
<td>1 (1.0)</td>
<td>7 (6.9)</td>
<td>0</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CRP mg/L, mean ± SD</td>
<td>11.0 ± 26.9</td>
<td>4.1 ± 7.3</td>
<td>14.0 ± 18.7</td>
<td>1.2 ± 1.8</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>Current NSAID treatment, n (%)</td>
<td>18 (25.0)</td>
<td>36 (35.0)</td>
<td>98 (96.1)</td>
<td>39 (62.9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Current PPI treatment, n (%)</td>
<td>6 (8.3)</td>
<td>-</td>
<td>40 (39.2)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Current systemic corticosteroid treatment, n (%)</td>
<td>26 (36.1)</td>
<td>17 (16.5)</td>
<td>5 (4.9)</td>
<td>0</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>Current csDMARD treatment, n (%)</td>
<td>33 (45.8)</td>
<td>6 (5.8)</td>
<td>3 (2.9)</td>
<td>-</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>Sulfasalazine, n (%)</td>
<td>1 (1.4)</td>
<td>1 (1.0)</td>
<td>2 (2.0)</td>
<td>-</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>Methotrexate, n (%)</td>
<td>0</td>
<td>5 (4.9)</td>
<td>1 (1.0)</td>
<td>-</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>Mesalazine, n (%)</td>
<td>10 (13.9)</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>Azathioprine, n (%)</td>
<td>21 (29.2)</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>Naive to csDMARD treatment, n (%)</td>
<td>23 (31.9)</td>
<td>90 (88.2)</td>
<td>98 (95.1)</td>
<td>0</td>
<td>&lt;0.001†</td>
</tr>
<tr>
<td>Current bDMARD treatment, n (%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Naive to bDMARD treatment, n (%)</td>
<td>70 (97.2)</td>
<td>95 (92.2)</td>
<td>81 (79.4)</td>
<td>-</td>
<td>&lt;0.001†</td>
</tr>
</tbody>
</table>

* Hyphens represent data that were not available or for which a statistical test was not meaningful.
‡ Calculated P values represent chi-square or Kruskal-Wallis tests between all (n = 4) groups.
† Calculated P values represent chi-square or Kruskal-Wallis tests between disease cohorts only.

High-level variation and taxonomic diversity of gut microbiota. We performed 16S rRNA sequencing and taxonomically profiled a total of 339 stool samples (277 patients and 62 disease-negative back pain controls). At the phylum level, our disease cohorts were dominated by Firmicutes, followed by Bacteroidetes, Actinobacteria, and Proteobacteria (Supplementary Figure 2A and C, available on the Arthritis & Rheumatology website at https://onlinelibrary.wiley.com/doi/10.1002/art.42658). At the genus level, we observed substantially more variation between individuals in all cohorts, with Bacteroides, Prevotella, and Faecalibacterium comprising the dominant genera in each cohort (Supplementary Figure 2B and D). Notably, back pain controls and patients from the axial SpA and AAU cohorts had more Prevotella-dominant individuals on average than did the CD cohort, which had the highest proportion of Bacteroides-dominant individuals. Patients with CD phenotypes had the lowest alpha diversities (Figure 1A and C), but beta diversity did not appear to correlate with disease phenotype (Figure 1B) or technical variation due to the sequencing experiment (Figure 1D). Our samples clustered into three discrete enterotypes (Supplementary Figure 2E and F, available at https://onlinelibrary.wiley.com/doi/10.1002/art.42658), which significantly differed from one another in terms of median alpha diversity (Figure 1E). In line with previous reports (25), we found the continuous Prevotella/Bacteroides ratios of our samples to correlate with our projected ordination axes (Pearson’s r = −0.8, P < 0.001), accounting for up to 10% of between-sample variation (Figure 1F), and indicative of an inverse relationship between these two taxa.

Taxonomic associations with immune-mediated inflammatory disease states. To uncover taxa potentially mediating the concomitance of SpA, CD, and AAU, we performed a differential abundance analysis with disease-negative controls. We observed significantly lower mean relative abundances of...
several Firmicutes in the disease group as a whole, mainly from ASVs belonging to the Lachnospiraceae family (Figure 2A). Of these, Fusicatenibacter (ASV12), Anaerostipes (ASV21), and Blautia (ASV30) were among the most prevalent taxa across all samples. One ASV from the Christensenellaceae family, a health-associated taxon with known heritability (26,27), displayed the strongest overall depletion in the GESPIC group relative to the control group; when all seven ASVs in our data belonging to this family were binned, the magnitude of the depletion increased (Figure 2B). Similar behavior was observed for Fusicatenibacter, Marvinbryantia, and Erysipelotrichaceae ASVs but not for Bacteroides, Blautia, or Lachnospiraceae ASVs, which were, in contrast, not significantly differentially abundant at higher taxonomic levels. GESPIC patients also had significantly higher abundances of Collinsella and Flavonifractor, both of which tracked closely with single ASVs and were ultimately found to be related to the SpA and CD phenotypes, respectively (Figure 2C).

In addition to Collinsella, the SpA group was enriched in several Lachnospiraceae ASVs, Subdoligranulum (ASV505), and Holdemana (ASV121) (Figure 2A), few of which were significant when binned (Figure 2B). Similarly, the CD group was strongly enriched in a Ruminococcus ASV and strongly depleted in a Blautia ASV (Figure 2A), neither of which were significant at their respective genus levels (Figure 2B). In contrast, Phascolarctobacterium and Lachnoclostridium reached significance when binned but not at the ASV level (in the SpA and CD groups, respectively). There were no ASVs or higher-level taxa significantly associated with the AAU phenotype and only two weakly enriched Lachnospiraceae ASVs in the SpA+AAU phenotype; however, adjusted mean differential abundances for these groups mostly followed the pattern of the other case-control comparisons (Figure 2A). Taken together, our findings broadly reflect the known taxonomic diversity of the Lachnospiraceae family at the clinical level (28,29) and reveal shared and disease-specific taxonomic signatures ranging from ASV to family resolution.

**Impact of drug therapies in GESPIC patient microbiota.** To clarify the disease-associated microbiota signals, we examined a wide range of clinically relevant factors.
The immune-mediated disease phenotypes share a depletion of Lachnospiraceae taxa and an enrichment in **Collinsella**, especially in spondyloarthritis (SpA). Disease groups consisted of patients without concomitant disease unless explicitly stated (e.g., SpA + acute anterior uveitis [AAU]).

**A** shows phylogenetic relationships, regression coefficients, and prevalence of ASVs significantly associated with one or more disease states. Bolded taxa had an adjusted significance level >0.001 (corresponding to a mean differential abundance between disease and control groups of 0.1%).

The x-axis broadly represents evolutionary distance based on the 16S V4 amplicons. **B** shows regression coefficients for the bolded taxa in **A**, plus any bins which were significantly differentially abundant, colored by genus or family with 95% bootstrapped confidence intervals (95% CIs) shown. Bolded taxa were significant using a blocked Wilcoxon test at a q < 0.05 cutoff. **C** shows a Venn diagram containing the bolded taxa from **B**.

**Figure 2.** Immune-mediated disease phenotypes share a depletion of Lachnospiraceae taxa and an enrichment in **Collinsella**, especially in spondyloarthritis (SpA). Disease groups consisted of patients without concomitant disease unless explicitly stated (e.g., SpA + acute anterior uveitis [AAU]).

**Methods**

Linear models were built for amplicon sequence variants (ASVs) and their genus- or family-level taxonomic bins (n = 123), containing an interaction term to account for technical variation (see Methods). The same groups were used for all three panels. Linear models were built for ASVs (n = 442) and their genus- or family-level taxonomic bins (n = 123), containing an interaction term to account for technical variation (see Methods).
found no significant associations between any taxa and age, sex, or body mass index, indicating our disease signals were unlikely to be confounded or mediated by demographic factors. Medication intake has been found to explain more variation in microbiome composition than disease status alone and is therefore an important factor to consider in association studies (16,17). In contrast to much of the literature (30–33), no individuals in our study were taking bDMARDs at the time of sampling, and a vast majority were naive to any biologic therapies, allowing us a better approximation of a true baseline disease signal. However, 70% (including back pain controls) were taking at least one glucocorticoid, conventional/synthetic (cs) DMARD, or NSAID, and 22% were taking two concurrently (Supplementary Table 3, available on the Arthritis & Rheumatology website at https://onlinelibrary.wiley.com/doi/10.1002/art.42658).

Because >95% of our patients with SpA were taking NSAIDs, we could not examine the disease signal independent of this effect; however, 60% of our control group was taking NSAID monotherapy, which was associated with 1% to 2% more

Figure 3. Fusicatenibacter is enriched in nonsteroidal antiinflammatory drug (NSAID) monotherapy, and Akkermansia is enriched in Crohn’s disease (CD) patients previously treated with conventional/synthetic disease-modifying antirheumatic drugs (csDMARDs). (A) Linear regression coefficients for disease-associated taxa (Figure 2B), estimating the differential abundance between medicated and unmedicated patients, either while adjusting for disease status and/or technical variation (two leftmost columns; see Methods) or resulting from a comparison within a specific subset of patients (specified underneath). Bolded taxa were most robustly associated with CD, spondyloarthritis (SpA), or the shared signal (Figure 2C). Because >95% of our patients with SpA were taking NSAIDs, we could not examine the disease signal independent of this effect; however, 60% of our control group was taking NSAID monotherapy, which was associated with 1% to 2% more

Fusicatenibacter (ASV12) and Subdoligranulum (ASV16) on average (Figure 3B). Interestingly, these two genera specifically were found to predict good response to csDMARDs in rheumatoid arthritis patients (34). 38% of our patients with SpA only were additionally receiving proton pump inhibitors, which was associated with an increase in Phascolarctobacterium (Figure 3B), likely confounding that signal (Figure 2B). No other disease-associated taxa appeared to be sensitive to the drugs we were able to test.

Nearly half our patients with CD only were on a form of csDMARD therapy at the time of sampling, which was associated with significantly fewer Prevotellaceae (Figure 3B). When comparing csDMARD-naive patients with CD to those that were currently or had previously received treatment, we found that treated individuals had increased Phascolarctobacterium and Akkermansia (Figure 3B). Although species-level identification from amplicon data is a contentious practice that we avoided here, it is worth noting that these individuals also had ~1% more of a Ruminococcus taxon (ASV17) identified as *R. gnavus* (28,35), enriched in our patients with CD and elsewhere (using metagenomic data) in
treatment-resistant patients with UC patients (36), as well as patients with SpA with concomitant CD (33).

**Potential mediation of inflammation and disease activity by HLA-B27 and microbiota.** Other clinically relevant covariates included the presence or absence of the HLA-B27 antigen, systemic inflammation (as measured by CRP), and disease activity scores. Because our disease phenotypes were significantly different in their CRP levels and expression of HLA-B27 (Figure 4A), we hypothesized that some of our shared disease-associated taxa or HLA- and inflammation-associated taxa might serve as potential mediators (see Methods). *Fusicatenibacter* was the taxon most strongly (negatively) correlated with serum CRP (Figure 4B), and we estimated about 19% of the increased CRP observed in our GESPIC group could be mediated by (the relative lack of) this taxon (Figure 4C). Several *Subdoligranulum, Roseburia, Lachnospira,* and *Faecalibacterium* ASVs as well as the total (binned) *Faecalibacterium* abundance were significantly higher in HLA-B27+ individuals independent of their disease status (Figure 4D). Regardless of group, none of the taxa we considered (or summary features such as taxonomic richness and alpha diversity) were implied to mediate the increased CRP observed in HLA-B27+ individuals (Figure 4E). However, in the group of HLA-B27+ patients with SpA only and HLA-B27− back pain controls, we estimated that *Faecalibacterium* (ASV4) could mediate ~15% of the (modestly) significantly increased disease activity in HLA-B27+ individuals with SpA (Figure 4F).

**Figure 4.** Decreased *Fusicatenibacter* may partially mediate increased C-reactive protein (CRP) in patients with spondyloarthropathy (SpA) and Crohn’s disease (CD), and increased *Faecalibacterium* may partially mediate higher disease activity in human leukocyte antigen (HLA)-B27+ patients with SpA. (A) Summary of Table 1 highlighting clinical factors differentiating our phenotypes. (B) Nonparametric effect sizes and false discovery rate (FDR)-corrected significance tests showing serum CRP correlations independent of disease status for all taxa across all individuals in our study. (C) Mediation analysis in which c represents the estimated total effect of disease presence on serum CRP, and the sum of both direct (c') and indirect (a × b) paths, with the estimates and significance of each component (a, b, c, c') illustrated (see Methods). When accounting for (decreased) *Fusicatenibacter* abundances, serum CRP estimates were (significantly) further increased, from 6.7 mg/l to 8.28 mg/l in the German Spondyloarthritis Inception Cohort (GESPIC) group. *q < 0.1, **q < 0.01, ***q < 0.001. (D) Similar to (B), with HLA-B27. (E) HLA-B27+ individuals (mostly SpA and acute anterior uveitis [AAU]) had increased inflammation and disease activity. (F) Mediation analysis in which c represents the estimated total effect of HLA-B27 expression status on Bath Ankylosing Spondylitis Activity Index (BASDAI) scores, which was greater when considering the enriched *Faecalibacterium* (ASV4) observed in HLA-B27+ patients with SpA only compared with HLA-B27− controls. csDMARD = conventional/synthetic disease-modifying antirheumatic drug; FDR = XXXX; NSAID = nonsteroidal antiinflammatory drug. Color figure can be viewed in the online issue, which is available at [http://onlinelibrary.wiley.com/doi/10.1002/art.42658/abstract](http://onlinelibrary.wiley.com/doi/10.1002/art.42658/abstract).
DISCUSSION

CD, AAU, and SpA share an established epidemiology, yet the pathophysiology underlying their concomitance remains unclear. We analyzed a large human cohort comprising all three diseases and deeply explored the taxonomic composition of the gut microbiota with clinical covariates and disease concomitance for the first time. Our results showed a shared depletion of predominately Lachnospiraceae taxa, most notably Fusicatenibacter, which partially mediated increased CRP and was most abundant in controls receiving NSAID monotherapy. Individuals with SpA had a robust enrichment of Collinsella relative to chronic back pain controls, and HLA-B27+ individuals (regardless of disease phenotype) displayed enriched Faecalibacterium.

It has been hypothesized elsewhere that HLA-B27 plays a causal role in SpA pathology (37) and that it shapes the gut microbiota composition (4). Our mediation analysis formalized this hypothesis and furthermore considered whether the increased Faecalibacterium (ASV4) we observed in HLA-B27+ patients with SpA could potentially mediate the increased BASDAI in those same patients. Our results imply that this bacterium contributes to discomfort, pain, and fatigue perceived in patients with SpA. This seems to contradict its known antiinflammatory role in IBDs (38) yet aligns with similar HLA-B27+− microbiota comparisons in the SpA literature (39) and our finding that abundances of this taxon were positively correlated with BASDAI scores (Spearman’s rho = 0.22, q = 0.009).

Although we did not observe a strong microbiota signal in our patients with AAU only, our results suggest that patients with AAU with concomitant SpA (92% HLA-B27+) are more similar to patients with SpA only than disease-negative controls (Figure 2). Association studies seeking to link the HLA-B27+ microbiota to an existing clinical understanding of these diseases (20,40) should prioritize inclusion of non-SpA HLA-B27+ groups for more precise, powered comparisons, both from other diseases (ie, AAU as we did here) and from disease-negative and genetically similar individuals, as in Berland et al (33). Our control group mirrored the European population and was mostly HLA-B27− but had chronic back pain. Because half were receiving NSAID monotherapy, we were able to estimate the impact of this treatment on individual taxa and further strengthen our hypothesis that Fusicatenibacter is a key microbe mediating host gut-joint inflammation. This genus has only one known species, F. saccharivorans (identified as such in our data as ASV12), which alleviated colitis in a murine model and induced antiinflammatory interleukin-10 (IL-10) production in lamina propria cells from patients with UC (41). Blautia emerged from our analyses as another genus likely to play a role in immune-mediated inflammatory pathologies (42,43), although one that should perhaps be characterized with more genomic resolution than we had access to here (28).

Restriction to amplicon sequencing data was a major limitation of our study, precluding further characterization important for understanding potential disease pathomechanisms. Yet, the ASVs analyzed here represent de novo sequences that more accurately capture taxonomic diversity and are more reproducible than operational taxonomic units (OTUs) (44). Furthermore, the functional microbiota is intrinsically coupled to a set of discrete bacterial units (taxa) in an ecological context (45,46), and although some functional differences indeed correlate with strain-level organisms or even nucleotides—as in probiotic therapies (47) and bacterial single nucleotide polymorphisms (48), respectively—higher-level bacterial taxonomies are still clinically useful to stratify disease patients and generate testable hypotheses in experimental or animal disease models.

For example, our patients with SpA exhibited higher abundances of Collinsella—in line with previous results (49,50)—a genus that has elsewhere been shown to reduce the expression of enterocyte tight junction proteins in vitro (potentially contributing to gut leakage in vivo), as well as increasing the production of proinflammatory IL-17A and transcription factor NFκB1 (46). This is relevant in SpA, in which an excessive activation of IL-17A drives an expansion of Th17 cells, which further perpetuates IL-17A production (and thereby contributes to chronic inflammation) (51). Similarly, colitis is associated with hyperproduction of Th17 cells, partly resulting from dysregulated NFκB activation responsible for inflammatory T cell differentiation (52). Previous CD-focused work isolated a 15-kDA microbial antiinflammatory molecule from F. prausnitzii able to inhibit NFκB signaling in vitro and alleviate colitis in mice (53); however, Faecalibacterium (ASV4) was not identified as this species in our data (although others in Figure 4D were). Like Blautia, Faecalibacterium taxa appear to require at least metagenomic resolution to unravel their clinically relevant properties (54).

Here, we presented the baseline cross-section of a prospective cohort examining SpA, CD, and AAU. Taken together, our results suggest that there is much more to be uncovered about the immunomodulatory properties of certain bacteria in these epidemiologically related pathologies, especially at the molecular level, to eventually leverage the diagnostic and therapeutic potential of the microbiome. Experimental work is needed to validate our findings, especially those that identified potential mediators of inflammation or disease activity. Future studies would benefit from whole (meta-)genome sequencing and fecal metabolite quantification (perhaps in parallel with serum) to better disentangle potential host and microbial contributions to inflammatory disease states (55). More end-to-end collaboration between clinicians, experimentalists, and statisticians is needed to design studies that integrate molecular omics approaches to understanding disease mechanisms (56–58) with established diagnostic and treatment criteria and biomarkers such as fecal calprotectin and serum zonulin (59,60).
AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Forslund had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Rios Rodriguez, Rademacher, Proft, Markó, Pleyer, Siegmund, Forslund, Poddubnyy.

Acquisition of data. Rios Rodriguez, Löber, Strowig, Marchand, Kirwan, Siegmund, Poddubnyy.

Analysis and interpretation of data. Essex, Forslund, Rios Rodriguez, Poddubnyy.

ACKNOWLEDGMENT

Open Access funding enabled and organized by Projekt DEAL.

REFERENCES

35. Henke MT, Kenny DJ, Cassilly CD, et al. Ruminococcus gravis, a member of the human gut microbiome associated with Crohn’s...


44. Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J 2017;11:2639–43.

