Alterations in the gut microbiome implicate key taxa and metabolic pathways across inflammatory arthritis phenotypes

Thompson, Kelsey N; Bonham, Kevin S; Ilott, Nicholas E; Britton, Graham J; Colmenero, Paula; Bullers, Samuel J; McIver, Lauren J; Ma, Siyuan; Nguyen, Long H; Filer, Andrew; Brough, India; Pearson, Claire; Moussa, Caroline; Kumar, Vinod; Lam, Lilian H; Jackson, Matthew A; Pawluk, April; Kiriakidis, Serafim; Taylor, Peter C; Wedderburn, Lucy R

Link to publication on Research at Birmingham portal

Alterations in the gut microbiome implicate key taxa and metabolic pathways across inflammatory arthritis phenotypes

One sentence summary: Gut microbiome configurations and activity exhibit similarities across distinct types of inflammatory arthritis.

Authors: Kelsey N Thompson[1,2,3], Kevin S Bonham[1,2], Nicholas E Ilott[4], Graham J Britton[5], Paula Colmenero[4], Samuel J Bullers[4], Lauren J McIver[1,3], Siyuan Ma[1,2], Long H Nguyen[1,3,6], Andrew Filer[7,8], India Brough[4,9], Claire Pearson[4], Caroline Moussa[4], Vinod Kumar[4], Lilian H Lam[4], , Matthew A Jackson[4], April Pawluk[1,3], The Inflammatory Arthritis Microbiome Consortium (IAMC) investigators group, Serafim Kiriakidis[9], Peter C Taylor[9], Lucy R Wedderburn[10,11,12], Brian Marsden[4], Stephen P Young[13], Dan R Littman[14], Jeremiah J Faith[5], Arthur G Pratt[15,16,17], Paul Bowness[9] , Karim Raza[7,8,13], Fiona Powrie[4], and Curtis Huttenhower[1,2,3,18]

Affiliations:

[1]Department of Biostatistics, T.H. Chan School of Public Health, Harvard University; Boston, MA, 02115, USA
[2]Infectious Disease and Microbiome Program, Broad Institute of MIT and Harvard; Cambridge, MA, 02142, USA
[3]Harvard Chan Microbiome in Public Health Center, Harvard T. H. Chan School of Public Health; Boston, MA, 02115, USA
[4]Kennedy Institute of Rheumatology, Nuffield Department of Orthopaedic, Rheumatology and Musculoskeletal Sciences, Oxford University; Oxford, OX3 7FY, UK
[5]Marc and Jennifer Lipschultz Precision Immunology Institute and Department of Genetics and Genomic Science, Icahn School of Medicine at Mount Sinai, New York, NY, 10029, USA
[6]Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School; Boston, MA, 02114, USA
[7]Institute of Inflammation and Ageing, College of Medical and Dental Sciences, University of Birmingham, Queen Elizabeth Hospital; Birmingham, B15 2TT, UK
[8]MRC Versus Arthritis Centre for Musculoskeletal Ageing Research and Research Into Inflammatory Arthritis Centre Versus Arthritis, University of Birmingham; Chesterfield S41 7TD, UK
[9]Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford; Oxford, OX3 7LD, UK
[10]Centre for Adolescent Rheumatology Versus Arthritis, University College London, UCLH, and GOSH; Chesterfield S41 7TD, UK
[11]NIHR Great Ormond Street Biomedical Research Centre, University College London; London WC1N 1EH, UK
[12]UCL GOS Institute of Child Health, University College London; London WC1N 1EH, UK
[13]Department of Rheumatology, Sandwell & West Birmingham NHS Trust; West Bromwich B71 4HJ, UK
[14]Howard Hughes Medical Institute and the Kimmel Center for Biology and Medicine of the Skirball Institute, New York University School of Medicine; New York, NY 10016, USA
[15]Translational and Clinical Research Institute, Newcastle University; Newcastle upon Tyne NE2 4HH, UK
[16]Research into Inflammatory Arthritis Centre Versus Arthritis, Newcastle Birmingham; Glasgow, and Oxford, Chesterfield S41 7TD, UK
[17]Department of Rheumatology, Musculoskeletal Services Directorate, Newcastle upon Tyne Hospitals NHS Trust; NE7 7DN, UK
[18]Department of Immunology and Infectious Diseases, Harvard T. H. Chan School of Public Health; Boston, MA, 02115, USA

1 # Abstract

2 Musculoskeletal diseases affect up to 20% of adults worldwide. The gut microbiome has been
3 implicated in inflammatory conditions, but large-scale metagenomic evaluations have not yet
4 traced the routes by which immunity in the gut affects inflammatory arthritis. To characterize the
5 community structure and associated functional processes driving gut microbial involvement in
6 arthritis, the Inflammatory Arthritis Microbiome Consortium investigated 440 stool shotgun
7 metagenomes comprising 221 adults diagnosed with rheumatoid arthritis, ankylosing spondylitis,
8 or psoriatic arthritis, and 219 healthy controls and individuals with joint pain without an underlying
9 inflammatory cause. Diagnosis explained about 2% of gut taxonomic variability, which is
10 comparable in magnitude to inflammatory bowel disease. We identified several candidate
11 microbes with differential carriage patterns in patients with elevated blood markers for
12 inflammation. Our results confirm and extend previous findings of increased carriage of typically
13 oral and inflammatory taxa, and decreased abundance and prevalence of typical gut clades,

1

14 indicating that distal inflammatory conditions, as well as local conditions, correspond to alterations
15 to the gut microbial composition. We identified several differentially encoded pathways in the gut
16 microbiome of patients with inflammatory arthritis, including changes in vitamin B salvage and
17 biosynthesis and enrichment of iron sequestration. Although several of these changes
18 characteristic of inflammation could have causal roles, we hypothesize that they are mainly
19 positive feedback responses to changes in host physiology and immune homeostasis. By
20 connecting taxonomic alternations to functional alterations, this work expands our understanding
21 of the shifts in the gut ecosystem that occur in response to systemic inflammation during arthritis.

Introduction

23 Alterations to the gut microbiome have been implicated in several inflammatory diseases,
24 particularly in the gastrointestinal tract, including the inflammatory bowel diseases (IBD) and
25 colorectal cancer (CRC)(*1, 2*). Although the role of gut microbes in other inflammatory conditions
26 such as type 1 diabetes (T1D) and metabolic syndrome have recently come under investigation(*3,*
27 *4*), their influences on or responses to systemic inflammation or disease progression remain
28 poorly elucidated. Inflammatory musculoskeletal arthropathies stand to benefit from a better
29 understanding of gut microbial ecology, both as an early biomarker for diagnosis of these
30 conditions, and as a potential new route for therapy. Arthropathies, including rheumatoid arthritis
31 (RA), ankylosing spondylitis (AS), and psoriatic arthritis (PsA), affect over 50 million adults
32 worldwide(*5-7*), who currently have no curative treatment options. Thus, understanding their
33 corresponding alterations within the gut microbiome is essential to both the underlying basic
34 biology driving systemic inflammation and clinical routes of arthritis treatment.

35 The etiology of many of the subtypes of arthritis can be traced back to aberrant immune
36 responses, which may be triggered or sustained by acute or long-term interactions with gut
37 microbial populations(*8*). This is true over and above human genetic contributions, which include
38 variants of the human leukocyte antigen (HLA) family(*9*). In RA, heritability is estimated at 60%,
39 although without clearly-resolved causal loci(*10-13*). Conversely, the heritability of AS
40 approached 90%, with HLA-B27 carriage the strongest genetic risk factor(*14-16*). Additionally,
41 smoking has been established as a likely trigger of RA, representing at least one specific
42 interaction between environmental and genetic factors in arthritis etiology(*13, 17*). The
43 microbiome is one of the most proximal forms of "environment", and indeed many arthritis risk
44 alleles such as HLA are, like those of IBD, known to be involved in microbial interactions or
45 immune sensing(*18*).

46 Since arthritis pathology is localized in the periphery, all of these arthropathies represent cases
47 in which any involvement of the gut microbiome would be "transmitted" systemically through
48 biochemical and immune-mediated signals. Research on this so-called "gut-joint-axis" dates back
49 to the 1890s, when researchers hypothesized that arthritic conditions could be caused by
50 *Mycobacterium* infections(*19*). Murine models have furthered this hypothesis by showing that
51 microbial disease triggers are required for SpA type arthritis to develop(*20-22*) and that gut
52 microbial colonization is necessary for Th17 differentiation (protecting germ-free mice from
53 disease)(*23*). Several strong indicators of the "gut-joint-axis" exist in humans as well, including
54 subsets of patients with chronic IBD exhibiting increased risk of peripheral arthritis(*24, 25*),
55 reactive arthritis occurring after pathogen infections(*26*), and the induction of autoreactive
56 cartilage degradation by specific bacterial strains(*25*). Several studies in smaller human
57 populations, primarily studied using 16S rRNA gene amplicon (16S) sequencing, found
58 compositional alterations of the gut microbiome in patients with RA, PsA, and AS(*27-41*). These
59 included the presence of clades that are frequently pathogenic, increased abundance of typically
60 oral microbes in the gut, and altered abundance of typical human gut clades(*39-42*). However,

61 there is no substantial agreement on which dysbioses are hallmarks of systemic inflammation in
62 arthritis. Additionally, 16S-based profiles do not provide direct insight into the functional
63 implications of microbial compositional changes, and thus far the agreement in functional changes
64 from the few shotgun studies is limited but have identified sweeping changes(*34, 36, 39-41, 43*).
65 Therefore, a comprehensive understanding of the role of the gut microbiome in arthritis
66 development and persistence is still lacking, which has the potential to better-support early
67 disease detection, prevention, or later-stage therapy.

68 Here, we introduce the work of the Inflammatory Arthritis Microbiome Consortium (IAMC), which
69 includes analysis of shotgun metagenome profiles spanning 440 participants with RA, AS, PsA,
70 and controls without inflammatory arthritis. We assessed the taxonomic and functional landscape
71 of the resulting gut microbiomes to elucidate key ecological and biochemical shifts linked to host
72 inflammatory responses and clinical arthritis phenotypes. In patients with inflammatory arthritis,
73 the overall compositional and functional profiles of the gut microbiome were substantially altered.
74 We identified enrichment of typically oral, pro-inflammatory, and mucin-degrading microbes, with
75 a corresponding decrease in several typical human gut-resident clades. Notably, several strains
76 of *Ruminococcus gnavus* isolated from human patients induced more severe phenotypes when
77 inoculated into mice. Further, several alterations in microbial community function were identified,
78 including the differential encoding of vitamin B salvage and biosynthesis and the encoding of folic
79 acid metabolism pathways. Similar to other local and distal inflammatory diseases, iron
80 scavenging was enriched in patients with current inflammation across heme, non-heme, and
81 siderophore-based mechanisms. Although these findings point to pathways and molecules of
82 interest and will serve as an important resource for hypothesis generation, future work will be
83 required to determine if these consistent functional changes occur causally, in response to
84 inflammatory arthritis, or both. At the least, our findings of community level taxonomic and
85 functional alterations in the gut microbiome implicate an interplay between host genetics, immune
86 system, and gut microbiome over the course of initiation, progression, and severity of arthritis.

Results

88 **Patient Cohort Characteristics**

89 We recruited 440 adults (ages 20 to 93) from different clinical locations in the United Kingdom,
90 Oxford (primarily AS patients), Birmingham (primarily RA patients), and Newcastle (primarily RA
91 patients), who met classification criteria for one of three arthritis subtypes or were included in the
92 non-inflammatory joint pain control group (**Fig. 1A-C**). Patient diagnoses included primarily
93 treatment naïve rheumatoid arthritis (RA, n=119), axial spondyloarthritis/ankylosing spondylitis
94 (AS, n=67), psoriatic arthritis (PsA, n=35), and non-inflammatory joint pain (typically fibromyalgia
95 [NIJP], n=54), as well as age-matched healthy controls (HC, n=165) (**Fig. 1A, table S1**). Although
96 some members of the NIJP category had low-titer autoantibodies for either rheumatoid factor (RF)
97 or anti-cyclic citrullinated peptide (anti-CCP), none were considered by consulting
98 rheumatologists to have clinically suspect arthralgia with respect to RA(*44*).

99 Earlier studies have identified alterations in gut microbial taxonomic profiles with both arthritis
100 diagnosis and local and distal inflammation(*27-33*). To expand these results, we focused on
101 patient diagnosis and current degree of disease activity as primary outcomes while adjusting for
102 relevant clinical covariates including patient age, current arthritis-related drug use, and disease
103 duration, as well as technical confounders such as sequencing batch and clinical site (**fig. S1A
104 to I**). Only a small subset of patients (58/440) reported use of antibiotics in the last six months,
105 ~half (22) within the previous two months and none at the time of sampling. These were equally
106 spread across all patients and HCs (**table S1**); such non-recent antibiotics use corresponded with

very little overall variation (PERMANOVA R^2=0.0049), and no individual features were associated with antibiotics use (MaAsLin 2). Antibiotic covariates were thus omitted from further analyses. We defined disease activity using two variables: (1) discretized C-reactive protein values (CRP) as a marker of current systemic inflammation (**Fig. 1D**), and (2) serum hemoglobin concentrations, as many patients with inflammatory arthritis also experience anemia as a feature of chronic disease (**fig. S1I**).
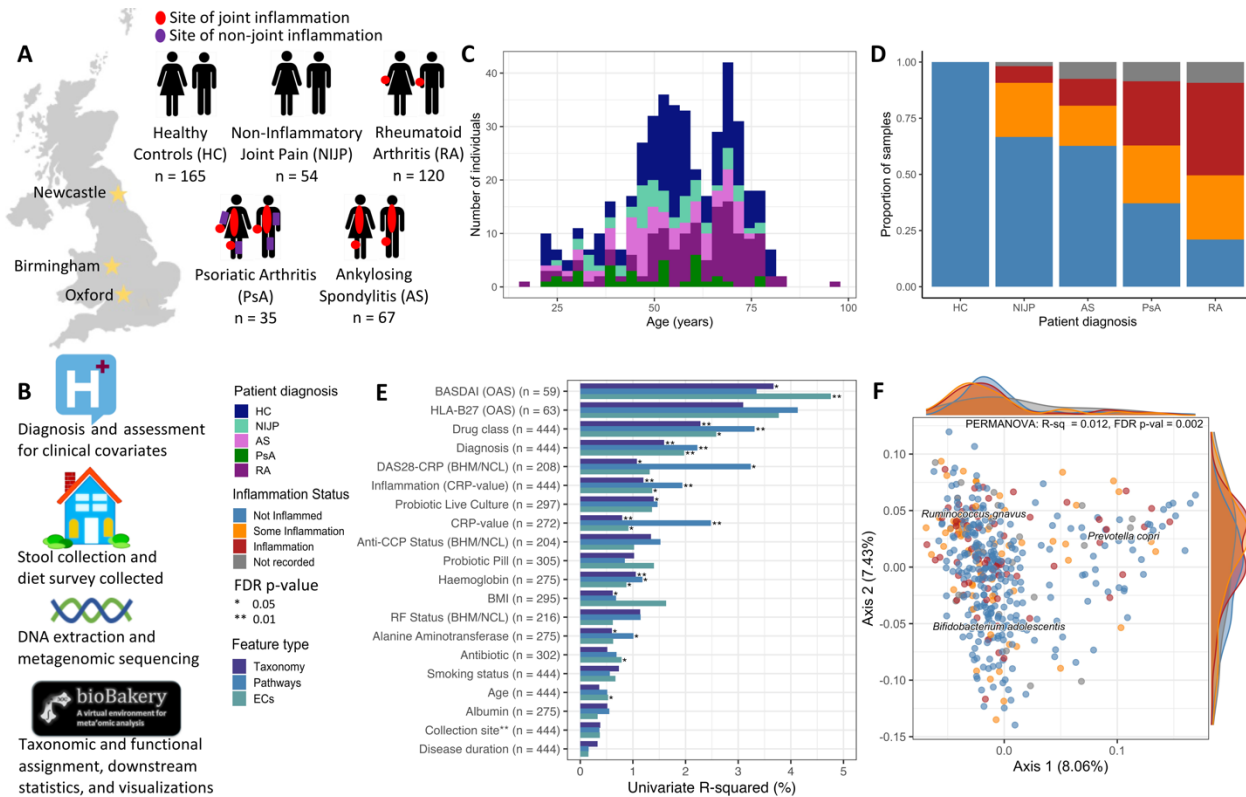
Figure 1: The gut microbiome is altered in patients with inflammatory arthritis. (A) Overview of the participants and samples collected from each subtype of arthritis and the **(B)** collection schema. **(C)** Prevalence of arthritis subtypes by age collected under this cross-sectional study by the Inflammatory Arthritis Microbiome Consortium (IAMC). **(D)** Proportion of patients with overt inflammation, some inflammation and no inflammation defined by tertiles of circulating serum concentrations of CRP by diagnosis (0 to 4mg/L, Not inflamed; 4 to 10mg/L Some inflammation; 10 to 167mg/L Inflammation). **(E)** Univariate PERMANOVA of Bray-Curtis dissimilarity by demographic and clinical measures. Color indicates the features assessed (microbial taxonomy, metagenomic pathways, and metagenomic Enzyme Commission (ECs) numbers). Tests are batch adjusted (**Methods**). **The effect of the collection site was also adjusted for by diagnosis to account for the fact that some sites only enrolled individuals with a particular disease subtype. **(F)** Bray-Curtis principal coordinates analysis of all 440 taxonomic profiles. Species are labeled using weighted averages and limited to just species whose abundance explains samples that fall outside of 0.04 distance from origin.

Individuals' disease activities varied from low to high disease activity [Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) = 0.08 to 8.4 (AS-specific measure); Disease Activity Score 28 for RA with CRP (DAS28-CRP) = 1.54 to 8.01 (RA/PsA-specific measure)] (**Fig. 1C**, **table S1, fig. S1C and D**). Although not ideal measures of inflammation, CRP concentrations for each participant provide a measure that was collected consistently across all population cohorts and ranged from 0 to 167 mg/L (**fig. S1C and D**). Thus, as the most consistently collected measure, we used CRP as a proxy for systemic inflammation; when available, it compared favorably to more direct measures such as BASDAI for AS or DAS28 for RA. Due to the asymmetric nature of these data, we categorized these patients using tertiles into three categories: not inflamed (0 to 4 mg/L), some inflammation (4 to 10 mg/L), and inflammation (greater than 10 mg/L). Healthy

4

135 controls only contributed fecal samples, not blood, and we could thus not quantify CRP
136 concentrations from these individuals and categorized all controls into the "not inflamed" group
137 **(Fig. 1D)**. From here on, we refer to the discretized CRP-value for systemic inflammation as
138 simply "inflammation." Anemia was also quantified in this population by current hemoglobin
139 concentrations, with anemia called when hemoglobin was less than 120g/L or 135g/L for females
140 and males, respectively (**fig. S1I**). Human leukocyte antigen B27 (HLA-B27) status was quantified
141 as either negative or positive, but only for the patients from Oxford (AS patients, n=67). In addition,
142 RF and anti-CCP status was categorized as negative or positive for RA patients (n=113) **(table
143 S1, fig. S1F to H**).

The human gut microbiome is altered in inflammatory arthritis.

145 Alterations in the overall composition of the gut microbiome were identified in patients with
146 inflammatory arthritis. Patient diagnosis explained a maximum of 1.6% and 2.3% of the
147 compositional differences in the taxonomic and functional microbial profiles, respectively, after
148 adjusting for the sequencing batch [Bray-Curtis PERMANOVA; false discovery rate (FDR) p-value
149 = 0.003 and 0.006]. In pairwise comparisons, these results were driven largely by differences in
150 the RA patients (**fig. S2**). Categorized CRP values, which represent the current amount of
151 inflammation a patient is experiencing, accounted for a maximum of 1.2% (FDR p-value = 0.003)
152 and 2.0% (FDR p-value = 0.006) of the variation in the composition of the taxonomic and
153 functional profiles, respectively (**Fig. 1E, fig. S2**). Inflammation thus explained a small but notable
154 shift in the overall gut microbial composition, not greatly below the amount often observed in
155 IBD(*1*) (**Fig. 1F**). Clinical measures of inflammation, such as the patients' DAS28-CRP and
156 BASDAI also explained similar amounts of variation within the gut ecology (**Fig. 1E**). This
157 indicates that systemic inflammation during arthritis, as characterized by either disease-specific
158 markers or circulating measures in all patients, corresponds with a substantial amount of variation
159 in the patients' gut microbiomes. Intriguingly, similar amounts of variation were also explained by
160 a patient's hemoglobin concentrations (g/L) (Taxonomy; R^2 = 1.1% and FDR p-value = 0.003,
161 Pathway; R^2 = 1.2% and FDR p-value = 0.009, **Fig. 1E**). Similar effect sizes also demonstrate a
162 consistent, but diverse, coupling of taxonomic and functional aspects of the gut microbiome, as
163 expected. HLA-B27, anti-CCP and RF status all did not induce alterations in the overall
164 composition of the gut microbiome (Bray-Curtis PERMANOVA taxonomy; FDR p-value >0.01).
165 Taken together, these results indicate that patients with inflammatory arthritis do harbor broadly
166 different configurations of microbes within their gut when compared to similarly-aged healthy
167 controls, consistent with previous studies(*28, 34, 42*).

168 HLA-B27 status explained a relatively large amount of variation in the composition of the gut
169 microbiome. However, none of these associations were significant after FDR-correction (Bray-
170 Curtis PERMANOVA taxonomy; FDR p-value = 0.605, pathways; FDR p-value = 0.381, ECs; FDR
171 p-value = 0.691), likely due to reduced sample numbers, as only the samples from patients with
172 AS and controls from Oxford had this information available (n = 135) **(Fig. 1E)**. Previous studies
173 have identified per-feature and overall compositional differences in the gut microbiome of patients
174 with HLA-B27(*45, 46*). Those genetic loci that do explain compositional shifts typically impact the
175 immune system, which in turn shapes (and is shaped by) the microbiome(*8*). Additionally, both
176 the anti-CCP antibody and the RF-status of the RA patients explained less than 2% of the variation
177 in the gut microbiomes and were not statistically significant (**Fig. 1E**). Taken together, these
178 results indicate that patients with inflammatory arthritis do harbor broadly different configurations
179 of microbes within their gut when compared to similarly-aged healthy controls, consistent with
180 previous studies(*28, 34, 42*).

Microbial taxonomic alterations occur in rheumatoid arthritis and ankylosing spondylitis.
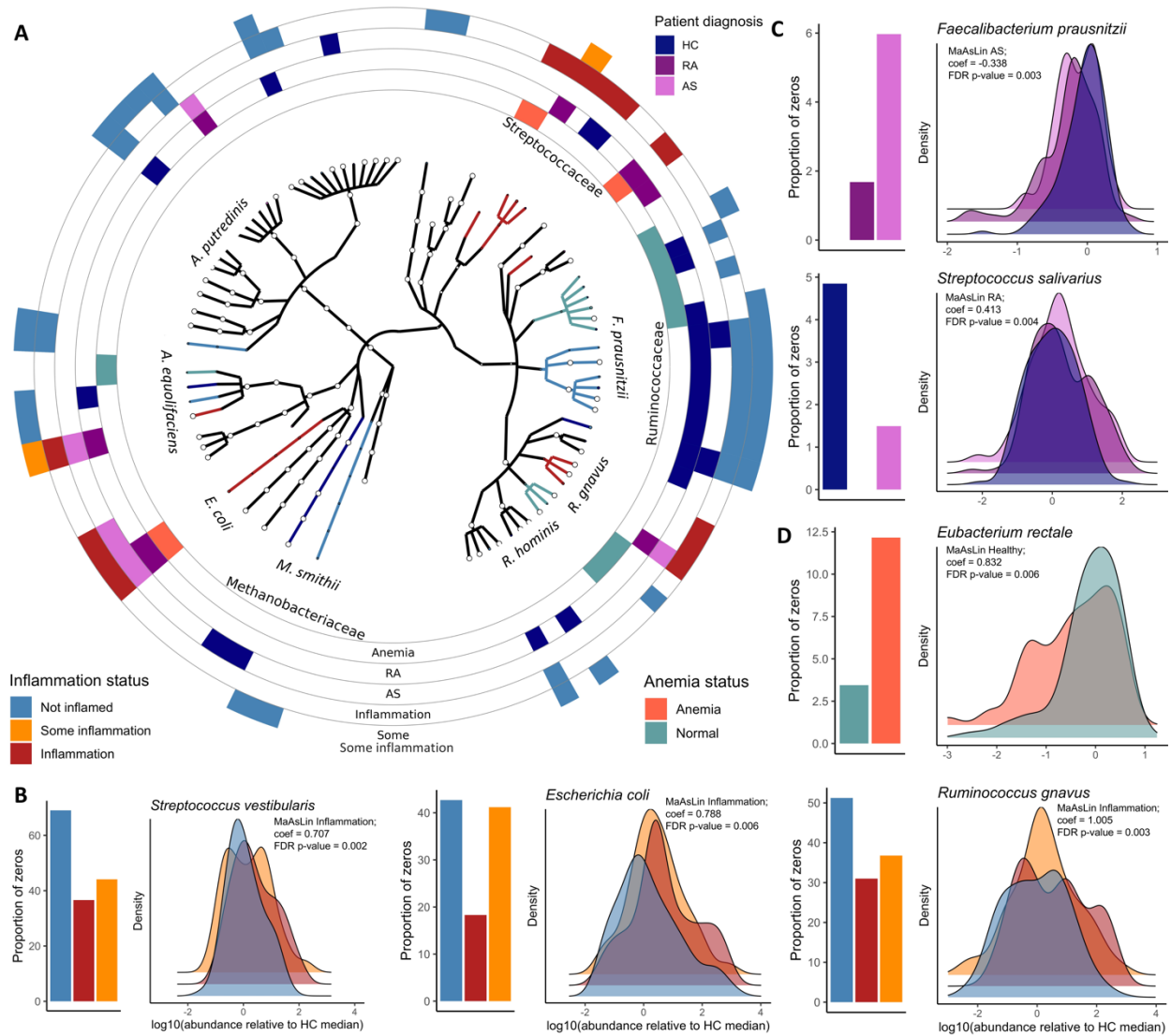
Figure 2: Taxonomic features differ in both a diagnosis- and severity-specific manner. (A) Shown are clade-specific taxonomic alterations of the gut microbiome by inflammation (inflamed versus not inflamed, some inflammation versus not inflamed), diagnosis (RA versus HC, AS versus HC), and hemoglobin concentrations (Anemia versus Normal). All associations were identified using MaAsLin 2. Internal branches are colored by strongest association, while outer rings summarize all associations. Leaf size was set as -log(pval)*sign(coef). These variables have been reduced for clarity; all results are presented in **fig. S4**. (B) Microbial species associated with inflammation status were determined based on serum CRP concentrations. Changes were observed in both the profiles of each clade's relative abundance within the gut community and their prevalence, here represented by the total number of zero abundance samples present for each condition (proportion of zeros). Increased abundance and prevalence were observed in three previously inflammation-associated species; *Streptococcus vestibularis, Escherichia coli,* and *Ruminococcus gnavus*. (C) Disease-specific alterations in the gut microbiomes were correlated with patients' inflammatory arthritis diagnoses. *Faecalibacterium prausnitzii* exhibits a lower abundance and prevalence in patients with RA, and to a lesser extent, AS, while *Streptococcus salivarius* had a higher prevalence and abundance especially in RA patients. (D) Several gut microbes exhibited strong correlations between several typical human gut residents and hemoglobin concentrations (d/L), with the highest effect size observed in *Eubacterium rectale.*

We identified several taxa associated with inflammation, diagnosis, and anemia or more disease-specific markers of inflammation (e.g. BASDAI or DAS28-CRP) that paralleled changes previously observed in dysbiotic individuals with IBD(*1*), including the clades *Streptococcus sp*., *Escherichia*

201   *coli*, and *Ruminococcus gnavus*(*47*) (**Fig. 2A, fig. S3 to 7**). Examining the prevalence of these
202   organisms across patients, it appears that *E. coli* and *R. gnavus* may exhibit a high abundance
203   but low prevalence phenotype, in which a small number of patients had substantially higher
204   abundances of these taxa. Previously, this pattern was observed with *Prevotella copri* in patients
205   with treatment-naïve RA (*29, 34, 48*), but that was not the case in this cohort (**fig. S8**). Several of
206   the clades that increased during inflammation are more commonly identified in the oral cavity(*49-*
207   *51*) than in the gut, including *Streptococcus mutans*, *S. vestibularis*, *S. salivarius*, and
208   *Bifidobacterium dentium* (**Fig. 2B and C, fig. S4, 5, and 7**). For these tests, we were careful to
209   adjust for proton pump inhibitor (PPI) usage, which has been hypothesized to facilitate the
210   transversion of oral taxa into the gut(*52, 53*). However, only 11 patients out of 275 were
211   documented to be actively taking PPIs at the time of sample collection, and we thus do not believe
212   this to be the mechanism. We do not have information on the oral health status (e.g. periodontal
213   disease) of these patients, although previous studies have found that the patients with RA were
214   four times as likely to have poor dental health(*54*). This strengthens the association of these
215   different types of microbial disruption, but leaves their respective causalities unclear.

216   Patients with IBD and T1D have both showed reduced abundance of the species
217   *Faecalibacterium prausnitzii, Roseburia intestinalis, Eubacterium rectale* and *Alistipes*
218   *putredinis*(*1, 55-57*), and we observed a similar decrease in our study of patients with AS and
219   RA, either by our proxy marker for inflammation or by disease phenotype or disease-specific
220   markers (**Fig. 2C and D, fig. S3, 4, and 7**)(*57-63*). In particular, *F. prausnitzii* and *R. intestinalis*
221   lost both abundance and prevalence in patients with current inflammation, whereas *E. rectale*
222   abundance was observed to be tightly coupled with hemoglobin concentrations. These microbes
223   are generally considered to be both highly responsive to inflammation and themselves anti-
224   inflammatory by routes such as short chain fatty acid (SCFA) production(*64-66*), making the
225   causality of these changes difficult to untangle observationally. However, it is striking that gut
226   microbial changes observed here for systemic inflammation during arthritis were both
227   taxonomically and functionally similar to those occurring during gastrointestinal inflammation,
228   which has been hypothesized to occur due to changes in oxygen availability in the gut
229   ecosystem(*67, 68*).
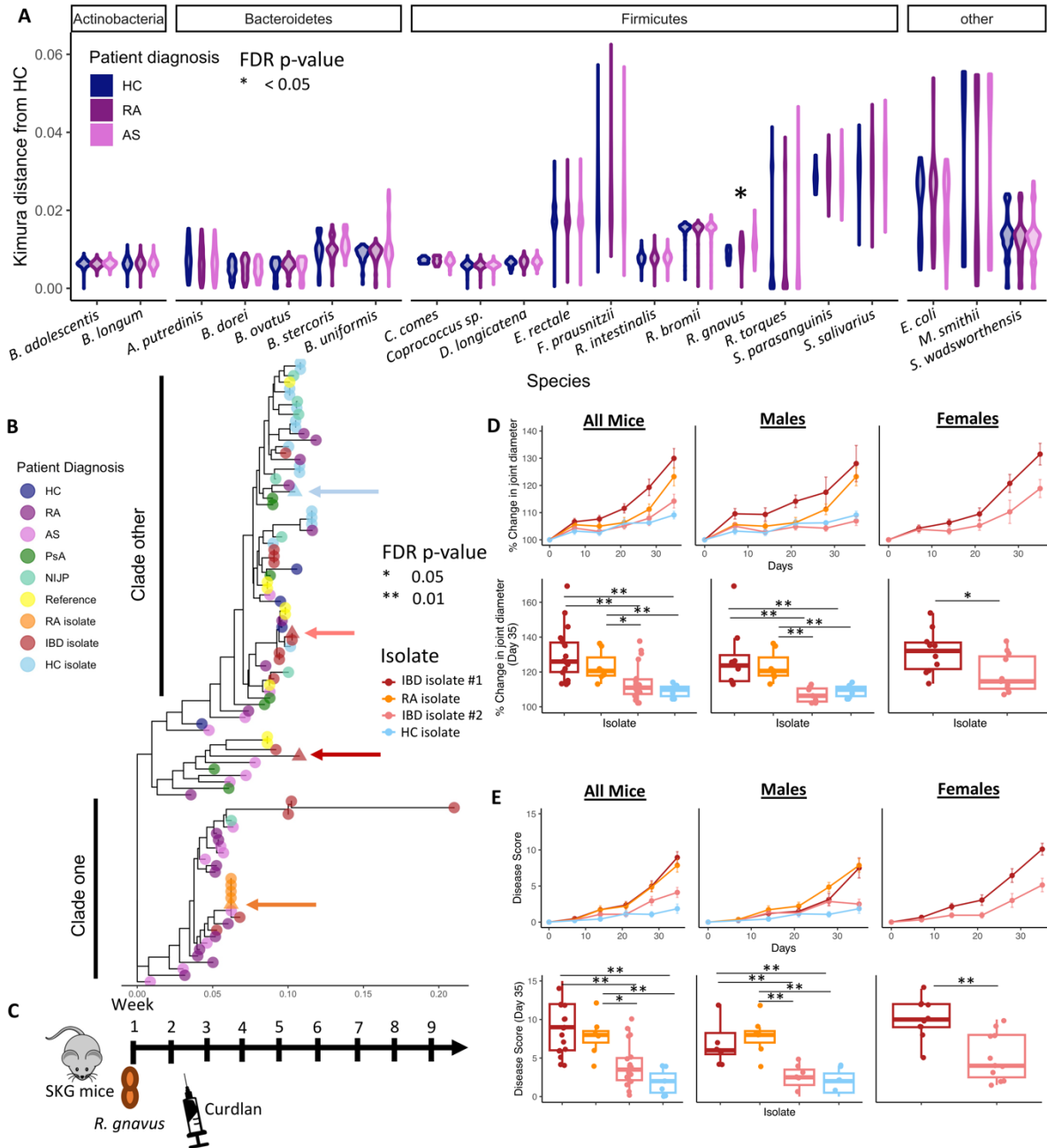
Figure 3: A *Ruminococcus gnavus* sub-species clade is enriched in patients with inflammatory arthritis. (A) Species were tested for sub-species phylogenetic structure associated with diagnosis and inflammation. Violin plots represent the spread and density of the pairwise 2-parameter Kimura distances between dominate strains within each sample compared only to the healthy control samples. Wilcoxon tests were used to determine significant changes in the pairwise distances between conditions and p-values were FDR corrected. **(B)** Phylogeny of *R. gnavus* strains from each individual's gut microbiomes. Isolates from NCBI and isolates cultured as part of this study (**Methods**) were included to add context to the subclade groupings. Triangles with arrows pointing at them indicate isolated strains used in (D) and (E). **(C)** Previously germ-free SKG mice were inoculated with a monoculture of a single *R. gnavus* strain (arrows on **Fig. 3B**) at week 1. Two to three weeks later curdlan was injected to simulate arthritis. **(D and E)** Male and female mice were then followed for joint diameter changes (**D**) and disease score (**E**). Longitudinal data are presented as mean disease score or joint diameter +/- the standard error. Box and whisker plots indicate the 25th, median, 75th and 1.5 times the interquartile range.

8

243 Within the ruminococci, a well characterized group of mucus-degrading bacteria(*69, 70*), *R. gnavus* has been implicated in many inflammatory conditions(*71-74*) and has been researched more extensively in IBD(*75-77*). Again, taxonomic associations with arthritis strikingly mimicked those of patients with IBD, though at a lower magnitude (**Fig. 2B; fig. S4-5, 7**). The abundance of *R. gnavus* was significantly increased in several patients with current high levels of CRP (Linear model; Not inflamed vs. Inflammation; coef = 1.005, FDR p-value = 0.003, **Fig. 2B, fig. S9-10**), interestingly including several NIJP subjects. Additionally, using both single nucleotide variants (SNVs, using StrainPhlAn(*78*)) and differences in pangenome-wide gene content (using PanPhlAn(*79*)), we identified phylogenetic structures that were significantly enriched in AS and RA patients (denoted Clade One; Kimura 2-parameter distance, PERMANOVA; R^2 = 0.18 , FDR p-value = 0.01, **Fig. 3A**) when compared the NIJP and HC individuals (which tended to carry members of Clade Other; **Fig. 3B** and **fig. S11 to 13**). To strengthen these results, an isolate from one RA patient also fell into Clade One, along with several isolates from a previously published IBD cohort(*75*). These results indicate that the presence of inflammation both locally and at distal locations in the host can correlate with structural, and potentially functional, changes in the gut microbiome. Other species tested did not exhibit the same subclade structuring as *R. gnavus* in this population (**fig. S9**).

260 Isolates from Clade One specifically enhanced inflammatory phenotypes when introduced into a mouse model of arthritis. New isolates (**Fig. 3B**) were derived from participant fecal samples and inoculated into previously germ-free SKG mice. Two to three weeks after the introduction of these monocultures of *R. gnavus*, curdlan was injected to induce arthritis symptomology. The presence of isolates from Clade One in the gut were able to potentiate the severity of arthritis-like symptoms in the SKG mouse, using both joint diameter and disease score as indicators of severity (**Fig. 3C to E**). This showed that the presence of these strains of *R. gnavus* is sufficient to induce a more severe phenotype, supporting its likely interaction with the immune system, as previously postulated(*76, 80*), and furthering arthritis symptoms.

Functional profiling reveals consistent functional alterations across all subtypes of arthritis.

271 We observed increased carriage of folate metabolism pathway and enzymes in individuals with arthritis and in those individuals with current high degrees of systemic inflammation (e.g. EC. 1.5.1.3 Dihydrofolate reductase; linear model not inflamed vs. inflammation; coef = 0.180, p-value = 0.028; **Fig. 4A to C, fig. S14 and 15**). However, carriage of the methionine cycle, which typically includes production of putrescine and homocysteine, was not consistently changed across both RA and inflammation (e.g. PWY-6151: S-adenosyl-L-methionine cycle I; linear model not inflamed vs. inflammation; coef = -0.03, FDR p-value = 0.17) (**Fig. 4A to C**). Alterations to nucleotide and amino acid pathways downstream of these processes are enumerated in **fig. S16** These lines of evidence suggest a dysregulation of folate metabolism in the gut ecosystem during arthritis, although this was distributed among a variety of different potential encoding organisms in different participants (**fig. S15**).
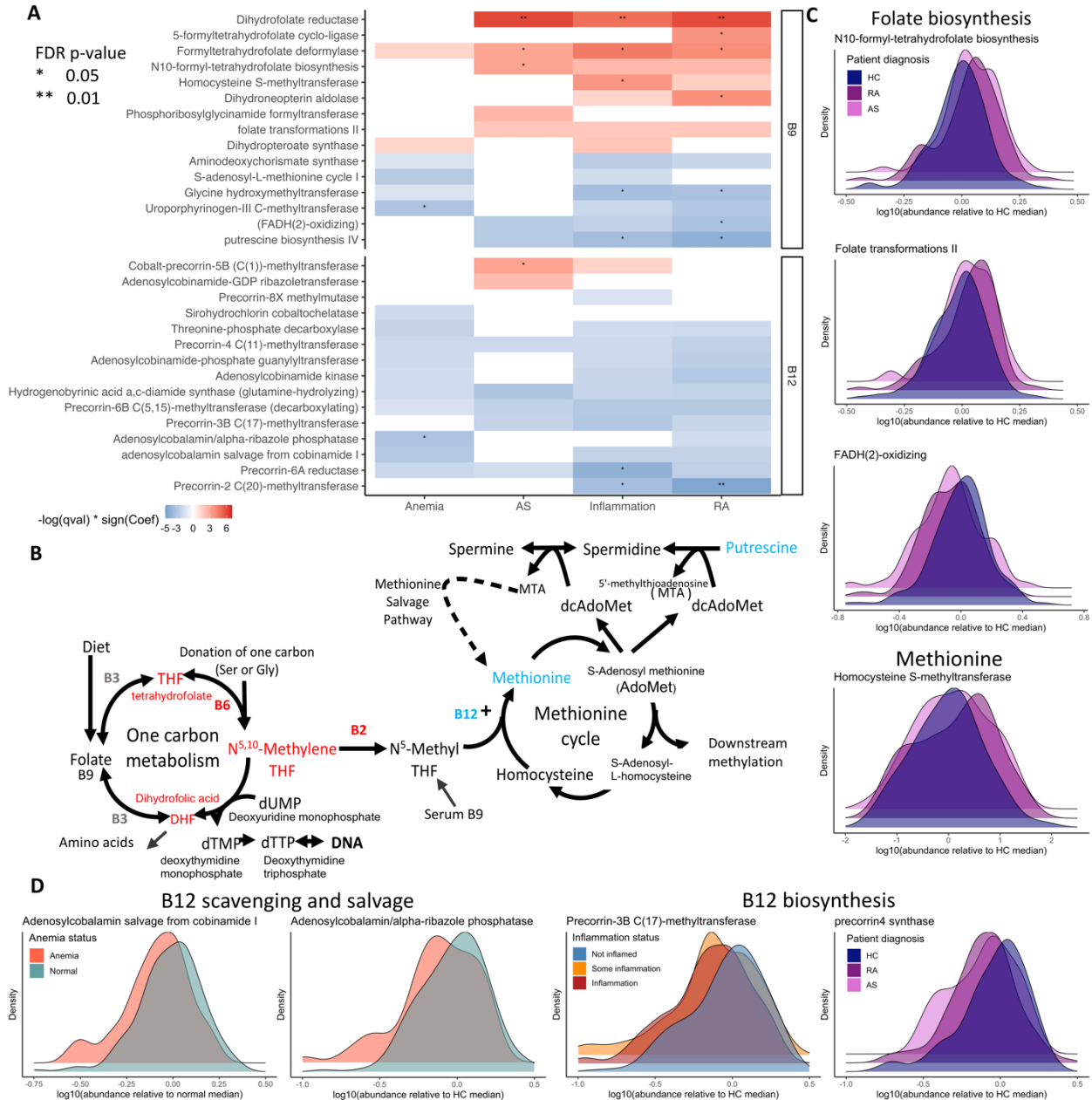
Figure 4: Differential encoding of vitamin B metabolism and processing were observed in the gut microbiomes of patients with inflammatory arthritis. (A) Shown are associations between patient diagnosis, inflammation status, and hemoglobin levels with folic acid metabolism, vitamin B12 salvage and biosynthesis and methionine biosynthesis and cycling. These associations were quantified using MaAsLin 2. **(B)** The folic acid metabolism pathway components carried by gut microbes. Enzymes in the one-carbon metabolism cycle that were enriched in gut metagenomes during inflammatory arthritis are shown in red, whereas enzymes decreased in the methionine cycle are shown in blue. **(C)** Encoding of folate metabolism cycle components within the gut microbiomes of patients was associated with a diagnosis of inflammatory arthritis. **(D)** Encoding of both the salvage and biosynthesis of vitamin B12 was correlated with anemia status, inflammation status, and patient diagnosis.

Patient hemoglobin concentrations and inflammation status were also associated with decreases in gut microbial carriage of vitamin B12 salvage and biosynthesis pathways (e.g. EC 1.2.1.54: Precorrin-6A reductase; linear model Normal vs. Anemia; coef = -0.17, FDR p-value = 0.14), in tandem with disruptions in vitamin B9 (folate) and its interconnected pathways (**Fig. 4A, B, and**

296   D). Vitamin B12 is a required cofactor in the link between the folate and methionine cycles(*81*)
297   (**Fig. 4B**). These alterations in the encoding of B12 salvage appear to be due to mainly (but not
298   entirely) to shifts in the relative abundance of *Eubacterium rectale* as noted above (**fig. S17**).
299   Additionally, the metagenomic abundances of several other enzymes associated with vitamin B
300   metabolism were also linked to patient diagnosis, inflammation status, or anemia status, including
301   vitamin B1 (thiamin/thiazole), vitamin B2 (flavin), vitamin B6 (pyridoxine), and vitamin B7 (biotin)
302   (**fig. S14, 18, and 19**).
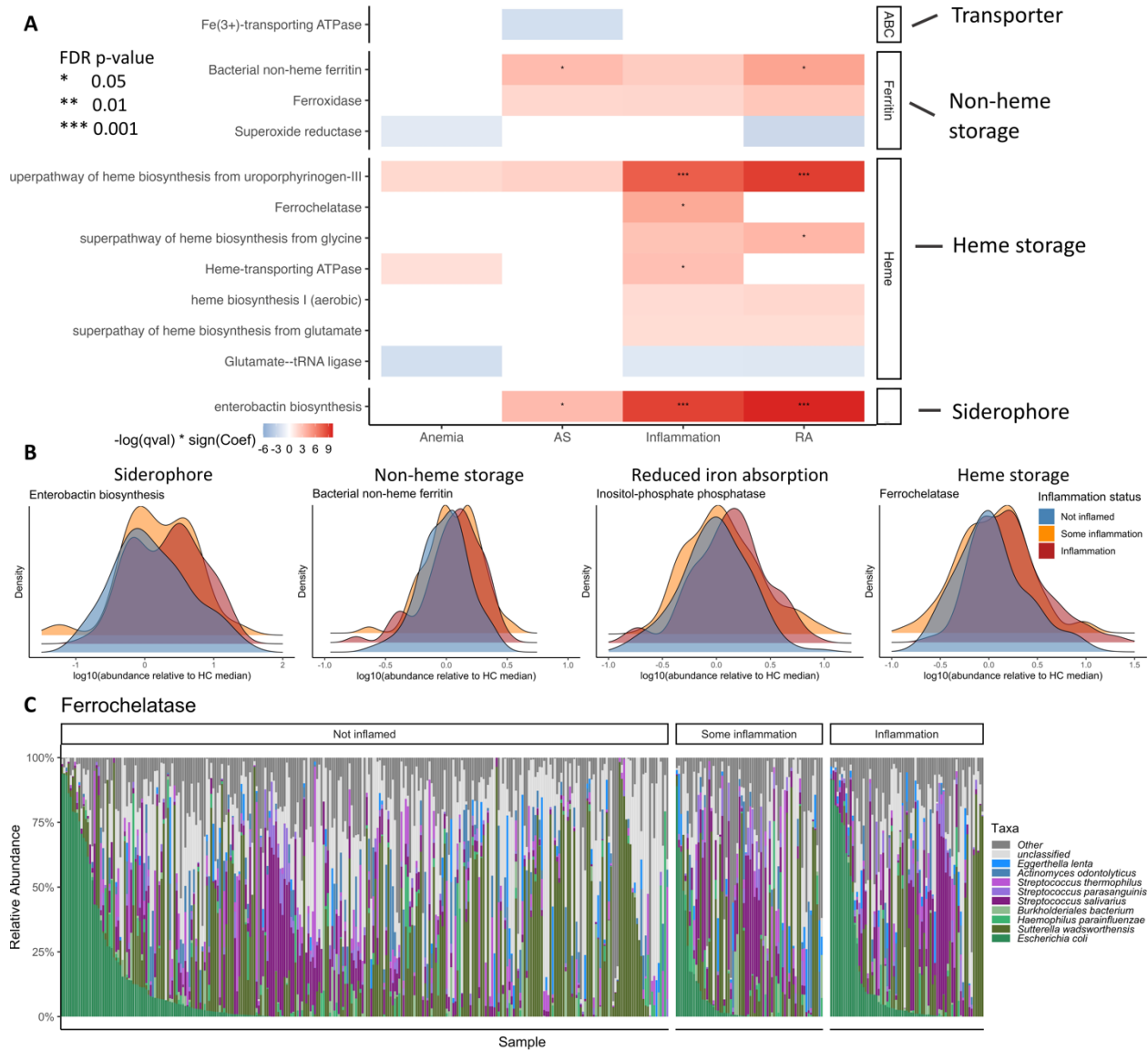
303

304   **Figure 5: Increased metagenomic carriage of pathways and enzymes involved in iron sequestration across**
305   **several distinct mechanisms were observed in individuals with arthritis with high serum CRP. (A)** Metagenomic
306   functional features (pathways and enzymes) associated with iron metabolism were correlated with the gut ecosystems
307   of patients AS, RA, evidence of anemia and elevated CRP values. Mechanisms of microbial iron sequestration included
308   heme and non-heme (ferritin) based storage, and high affinity siderophores. These associations were quantified using
309   MaAsLin 2. **(B)** Enzymes from several of the sequestration mechanisms were correlated with inflammation status. **(C)**
310   Species that most contribute to carriage of the ferrochelatase iron sequestration gene family are shown based on
311   inflammation status.

11

312 Alterations of other gut metagenomic pathways regulating metabolic cofactors were also identified
313 among inflammatory arthritis patients. Genes encoding several enzymes linked with microbial iron
314 sequestration were enriched in inflamed patients, as well as in RA and AS patients specifically
315 (e.g. ENTBACSYN-PWY: Enterobactin biosynthesis; linear model not inflamed vs. inflammation;
316 coef = 0.62, p-value = 0.00014). These included genes encoding ferrochelatase, bacterial non-
317 heme ferritin, ferroxidase, and heme biosynthesis (at the pathway level) (**Fig. 5A and B**). This
318 agrees with results in an earlier, smaller RA cohort in which oral and gut capacity for iron transport
319 was disrupted(*34*). However, this was previously attributed specifically to *Klebsiella* spp.; in
320 contrast, as with disruptions in folate metabolism, we found contributions to iron sequestration to
321 be encoded by diverse taxa (**Fig. 5C** and **fig. S20**), and for non-heme mechanisms, no one clade
322 was individually associated with their differential carriage. This again indicates that dysbiosis of
323 these processes can be distributed among different microbes in different people, or that there are
324 strain-level differential carriage within taxa, such as with ABC transporters within *R. gnavus* (**Fig.**
325 **3C**). However, heme-related mechanisms appear to be driven mainly by the high abundance
326 phenotype described above for *E.coli* in this population, as it is a major contributor to these
327 functions (**Fig. 5C** and **fig. S20**).

328 Several other functional classes, including both pathways and enzymes, exhibited differential
329 metagenomic carriage either by inflammation status or patient diagnosis (**fig. S21 to 25**). The
330 gene classes most highly associated with both RA and inflammation were those relating to the
331 production of isoprenoids or volatile hydrocarbons (VOCs) (**fig. S21**). Microbial gene families that
332 explicitly interact with host immunity were also differentially carried during arthritis, such as cell
333 wall remodeling proteins and oxidative stress response (**fig. S22 to 25**). Finally, enzymes involved
334 in the oxidative stress response, including methanogenesis, glutathione, and peroxiredoxin (**fig.**
335 **S22**) were differentially carried in inflammation. Also of current interest to the field, enzymes
336 involved in SCFA metabolism (*82*), 3-hydroxybutyryl-CoA dehydrogenases and short-chain acyl-
337 CoA dehydrogenase enzymes were also found to be differentially carried by the gut ecosystem
338 in inflammation (**fig. S23**). Overall, alterations to the functional landscape of the gut ecosystem
339 indicated disruptions in several key metabolism pathways during inflammatory arthritis.

340 Several other functional classes, including both pathways and enzymes, exhibited differential
341 metagenomic carriage either by inflammation status or patient diagnosis (**fig. S21-25**). The gene
342 classes most highly associated with both RA and inflammation were those relating to the
343 production of isoprenoids or volatile hydrocarbons (VOCs). Interestingly, the production of VOCs
344 under inflammatory conditions has been previously noted(*83, 84*), but not the involvement of the
345 gut microbiome in its upregulation. Several studies have also indicated anti-inflammatory
346 properties of isoprenoids, especially geraniol, farnesol, and geranylgeraniol(*85*). Here, we
347 observed that many gut microbial pathways with greater carriage during arthritis were involved in
348 geranylgeraniol biosynthesis (**fig. S21**). Thus, microbes within the gut ecosystem could be
349 increasing production of these small molecules if they are less bioavailable from the host or diet.

350 Microbial gene families that explicitly interact with host immunity were also differentially carried
351 during arthritis, such as cell wall remodeling proteins and oxidative stress response (**fig. S22-25**).
352 This trend was observed in relatively few genes and was especially true for patients with RA and
353 not those with AS. Enzymes involved in the oxidative stress response including methanogenesis,
354 glutathione, and peroxiredoxin were differentially abundant in these patients' gut metagenomes
355 (**fig. S22**). Finally, and of interest to the current short-chain fatty acid literature(*82*), we observed
356 a few select genes involved in butyrate metabolism to be differential, with the majority less
357 abundant in patients with RA or with higher circulating markers of inflammation. These included
358 the 3-hydroxybutyryl-CoA dehydrogenases and short-chain acyl-CoA dehydrogenase enzymes
359 (**fig. S23**). Overall, alterations to the functional landscape of the gut ecosystem indicated
360 disruptions in several key metabolism pathways during inflammatory arthritis.

Discussion

Here, we present the findings of a large cross-sectional study of adults with inflammatory arthritis diagnosis (and control participants), investigating alterations in gut microbiome composition and function associated with disease status and inflammation. The signals we detected associated with this family of systemic inflammatory conditions largely paralleled those identified in diseases defined by gastrointestinal inflammation, such as IBD. Changes in microbial taxa, functions (pathways and individual gene families), and in some cases even strains (such as those within the species R*uminococcus gnavus*) were shared between arthritis patients and other inflammatory diseases such as IBD, T1D, and other metabolic disorders. These changes were largely consistent among individuals with RA, AS, and PsA. In addition to the initial results presented here, the corresponding large shotgun metagenomic and clinical dataset offers the ability for further hypothesis generation and testing, including the potential for identification of additional arthritis therapeutic targets.

In particular, several previously-suggested "pro-inflammatory" microbes were enriched here during arthritis(*1*). This was particularly true for *E. coli*, which had an especially unique influence on the corresponding community functional potential, and has been previously shown to be enriched in many conditions including RA(*43, 86*). Mucin-degrading microbes such as *R. gnavus*(*68, 69*) were also differentially carried and functional during disease, down to the subspecies level. Recently, a substantial number of studies have found direct associations between *R. gnavus* and inflammation, including in arthritis(*84, 85*). Potential mechanisms include direct interaction of *R. gnavus* with the host immune system through extracellular proteins(*86*). Notably, a subset of phylogenetically distinct *R. gnavus* isolates from an individual in this cohort with RA and isolates from a healthy control and two IBD individuals were sufficient to increase arthritis severity when introduced into SKG, both supporting their causality and agreeing with previous subclade results. Finally, the presence of characteristically oral taxa in the gut microbiome of patients with chronic inflammation has been well documented, including in patients with IBD, UC, CRC, and metabolic disorders(*1, 48, 87, 88*). A few studies on patients with both AS and RA have also identified increases in streptococci in the gut(*21, 42*), similar to what was observed in our population.

Relatedly, *Prevotella copri* has been implicated as a potential disease trigger in RA, both epidemiologically(*35*) and in studies linking the HLA-DR-presenting peptide of certain strains of *P. copri* and the stimulation of a Th1 response in the onset of RA(*47*). However, across the microbiome studies conducted in primarily treatment naïve arthritis patients, there are conflicting reports regarding its role: some studies indicating increased burden of *P. copri* in the gut microbiomes of RA patients(*35, 47*), while others have found no link(*42, 61*). Additionally, there is evidence suggesting that treatment of RA patients with methotrexate (MTX) may revert the *P. copri* abundance back to normal levels(*89, 90*). In the current UK based study, we did not observe any evidence for increased *P. copri* abundance or prevalence in this cohort (**Fig. 1F** and **Fig. S8**). The RA patients included in this study were all disease modifying anti-rheumatic drug (DMARD) naive, thus none had yet been exposed to MTX or any other DMARD at the time of sample collection. Other studies with shotgun sequencing have also found other *Prevotella* species with increased abundance, which we also did not observe in our population(*39*). *P. copri* carriage has been shown to differ by both country of origin and diet(*91*), which could explain some of these differences.

These examples represent two ways in which our results generally agree with previous studies of the gut microbiome in inflammatory arthritis(*41, 43, 86*). We also observed similar broad patterns, such as the loss of typical gut consortia and increasing abundance of oral taxa and clades associated with gastrointestinal inflammation(*34-41, 43*). Although we were slightly hindered by

409 differences in collection targets within our disease subtypes (e.g. individuals with treatment-naive
410 early RA versus individuals with AS with predominantly controlled disease, as well as the larger
411 sample size in our RA group), microbiome alterations across different inflammatory arthritis
412 phenotypes were, when detectable, largely shared among such subpopulations. We found many
413 of the same microbes associated with either AS or BASDAI as we identified within our RA
414 individuals. We hypothesize this is most likely due to the consistent collection, sequencing, and
415 analysis methods applied throughout our cohort, as well as its relative geographical and
416 environmental homogeneity, any of which can otherwise cause inter-study differences(*87, 88*).

417 The observation of consistent shifts in the functional capacity of gut microbial communities in
418 patients with inflammatory arthritis provides the opportunity to explain their potential chemical and
419 regulatory consequences. These include changes in folic acid metabolism, iron sequestration,
420 metabolism of broad classes of B vitamins, and production of isoprenoids. Folic acid metabolism
421 in particular (microbial processing of folate to downstream compounds) was more abundant in
422 arthritis patients with higher circulating CRP. Methotrexate (MTX) is a dihydrofolate antagonist,
423 which competitively binds to and blocks several folate pathway enzymes. In patients with
424 treatment-naïve RA, higher basal folate metabolism has been documented in the peripheral
425 serum, and MTX treatment was shown to normalize that degree of folate metabolism(*89*).
426 Treatment with MTX often relieves arthritis patients of many of their joint inflammation symptoms,
427 indicating a potential role of folic acid metabolism in the disease etiology (although it is not clear
428 that this is the mechanism of action for MTX in this case). Almost no patients in this study were
429 currently taking MTX (specifically none of the RA patients), and increased microbial folic acid
430 metabolism thus appears independent of MTX exposure. Further, several studies have implicated
431 the role of well-regulated folate metabolism in the appropriate functioning of the host's immune
432 system(*90-93*), including natural killer (NK) cells(*90*), the proliferation of CD8$^+$ T lymphocytes(*92*),
433 the survival of FOXP3$^+$ regulatory T cells(*93*). One previous study in a smaller cohort has also
434 identified changes in folic acid metabolism pathways associated with disease improvement within
435 the gut ecosystem of patients with RA(*41*). Thus, folic acid metabolism within the gut microbiome
436 is a potential player in the etiology of arthritis, and warrants further mechanistic validation both
437 linked to and independently of MTX usage.

438 In an even clearer example of this causal vs. responsive dichotomy, increased carriage of
439 microbial iron sequestration via non-heme, heme, and ferroxide related mechanisms (**Fig. 5**)
440 could occur due to i) changes in the gut environment during disease that favor microbes
441 sequestering iron, or ii) greater microbial sequestration of iron as a contributing risk factor in
442 disease (or both). Notably, many different clades encoding iron sequestration systems were
443 enriched during arthritis, with no one primary driver taxon. Potentially relatedly, in RA, the immune
444 system has been shown to sequester iron away from other cell types, often resulting in
445 anemia(*94*). Previous studies in murine models have indicated that in response to iron, GIT
446 microbes are capable of both secreting small molecules that inhibit the transcription of HIF-2α,
447 which is responsible for the uptake of iron in the intestines, and concurrently upregulating their
448 own iron sequestration mechanisms resulting in decreased iron absorption in the host(*95, 96*). A
449 similar enrichment of iron sequestration genes was observed in patients with IBD, although with
450 a clearer corresponding hypothesis that it may be due to increased presence of blood within the
451 GIT(*97*). Even if true in IBD, this is unlikely to be the case in arthritis, where increased microbial
452 iron sequestration might instead result from anemic conditions within the host(*98, 99*).

453 Our study has limitations, despite increasing both the sample size and depth of microbial data
454 compared to previous studies, the inter-individual diversity of the human gut microbiome means
455 that our results are still derived from a relatively small sample size - notably from a single country
456 and dominated by a single ethnicity. This is especially true with the confounding nature of clinical
457 data, including site specific collection of distinct diagnoses, a large age range, inherent differences

458 in the sex distribution and uneven loading of arthritis subtypes across sequencing batches.
459 Further, since we only used sequencing data, especially since these data are based on DNA
460 profiles only, we do not have a true functional profile. Thus, as noted several times above, it is
461 impossible to establish the causality or mechanism of these gut microbial changes from an
462 observational human study, and we fully expect our own and others' longitudinal human and
463 model system research to clarify these.

464 However, this study comprehensively evaluates functional changes within the gut microbiome of
465 patients with RA and AS at scale. We found what are becoming canonical shifts in the distribution
466 of several microbial processes in the gut during inflammation, including for both local
467 gastrointestinal conditions and systemic inflammatory disease. Our study contributes to the
468 growing body of evidence that the gut microbiome and inflammation throughout the body are
469 tightly coupled, likely both casually and responsively, as the gut microbiome serves as a mediator
470 of environmental triggers and then also changes in response to immune activity. We hypothesize
471 that this occurs in part due to a functional "echo" of systemic inflammation in the gut microbiome,
472 due to the similarity in the specific processes that are altered in IBD and in arthritis. Some of these
473 alterations, such as those for B vitamin metabolism, including both B9 and B12, could represent
474 mechanisms for long-term prevention, risk reduction, or treatment, as could microbial iron
475 sequestration during arthritis-linked anemia. We thus expect these results and resources to
476 represent the next step in understanding and managing inflammatory arthritis through its interplay
477 with the gut microbiome.

Methods

478

Study Design

479

480 Participants were recruited for this multi-center study in Birmingham, UK (primarily RA and HC
481 patients, exact numbers in **table S1**), Newcastle, UK (RA, PsA, NIJP, AS, and HC), and Oxford,
482 UK (AS and HC only) from June 2015 until March 2020, samples were accepted until the last year
483 of the grant period. Patients enrolled for this study were aged 17 to 97 years. As expected, based
484 on disease epidemiology(*100, 101*), diagnoses were skewed by female sex, comprising 63%,
485 30%, 40%, 85%, and 58% of the patients with RA, AS, PsA, NIJP, and HC, respectively. The
486 majority of participants reported non-Hispanic white ethnicity (74%). Approximately 50% of our
487 population had never smoked cigarettes, and this was generally lower among cases than controls
488 (**table S1**). A simple power calculation based on the human microbiome project data(*102*),
489 indicated that at 400 samples (150 healthy controls and 250 cases and adjusting for 15 covariates,
490 we were well powered to quantify changes in both abundant and rare taxa, power = 1).

491 This observational cross-sectional study was designed as a sub-project within the Inflammatory
492 Arthritis Microbiome Consortium (IAMC). All samples were collected under the IAMC umbrella
493 biospecimen protocol from one of the three main collection sites. For this study we focused on
494 adults with clearly defined arthritis (corresponding to collection sites in Oxford, Birmingham, and
495 Newcastle, UK) to robustly establish the associations between inflammatory arthritis and the gut
496 microbial composition. Patients were consented by their treating physician to have blood and stool
497 collected. Other important clinical metadata captured by the treating physician included disease-
498 specific measures of BASDAI (for AS patients), DAS28 (for RA and PsA), and swollen/active joint
499 counts. All clinical and demographic information was curated and securely housed in
500 REDCap(*103, 104*). Participants provided written informed consent. The study was approved by
501 the Newcastle and North Tyneside Regional (REC 12/NE/0251), Oxfordshire (REC
502 06/Q1606/139) and West Midlands-Back Country (REC 12/WM/0258) Research Ethics
503 Committees.

504 Biological material was obtained from consenting patients referred from primary care with
505 suspected arthritis and seen in either the Newcastle Early Arthritis Clinic (NEAC), UK [Newcastle
506 upon Tyne NHS Foundation Trust], the Birmingham Early Arthritis Cohort from Sandwell and West
507 Birmingham NHS Trust, University Hospitals Birmingham NHS Foundation Trust, and the Nuffield
508 Orthopaedic Center (NOC) and Oxford biobank during 2017-2019. The recruitment strategy was
509 designed to minimize enrollment of individuals exposed to systemic corticosteroids or disease-
510 modifying anti-rheumatic drugs (DMARDs) prior to biological sampling. Clinical diagnoses were
511 ascertained by board-certified rheumatologists in accordance with standard and appropriate
512 classification criteria where available. RA was assigned only where 1987 American College of
513 Rheumatology or 2010 European League Against Rheumatism/American College of
514 Rheumatology classification criteria were fulfilled. All axial spondyloarthritis patients met the
515 Assessment of Spondyloarthritis International Society (ASAS) criteria for axial
516 spondyloarthritis(*105*). When plain radiographs were performed, the vast majority (>90%)
517 additionally met modified New York Criteria for Ankylosing Spondylitis and we have hence used
518 AS. A diagnosis of non-inflammatory joint pain (NIJP) was assigned when the consulting
519 rheumatologist considered the presentation neither attributable to an inflammatory arthritis, nor to
520 osteoarthritis. For the HC's from the Oxford biobank, samples were selected to enrich for HLA-
521 B27 positive individuals, specifically selecting for about 50% of the controls being positive for the
522 HLA-B27 allele.

Sample collection
524 Blood samples were collected at routine clinical visits. Whole blood was drawn and stored at -
525 80°C. For serum, blood was drawn into a 5-ml SST tube and allowed to clot at room temperature
526 after centrifugation for 15 min at 1000G. Serum supernatants were aliquoted and stored at -80°C
527 and in accordance with approved protocols. Blood samples were assayed for CRP, RF, anti-CCP,
528 HLA-B27 genotype, full blood count, and liver function by UK certified labs. RF and anti-CCP were
529 classified as positive or negative according to local laboratory cut-offs.

530 Those enrolled were provided with a previously-validated stool collection kit designed to maximize
531 ease of participation and to impart a minimum perturbative effect on downstream extraction and
532 computational protocols(*106*). They were also furnished with a brief dietary inventory modeled
533 after prior investigations and a questionnaire surveying various microbiome-relevant exposures,
534 such as the recent use of antibiotics, each completed at the time of collection. All EtOH-fixed kits
535 used for metagenomic sequencing were returned within 1 to 3 days of a matched blood sample
536 by Royal Mail Response Service Delivery. Participants used a FecesCatcher (Tag Hemi) to collect
537 a stool aliquot into a Sarstedt Feces container containing 100% molecular biology grade ethanol
538 (Merck Life Sciences) to facilitate preservation at more ambient temperatures. Participants
539 returned the samples enclosed in UN3373 Category B Postal kit boxes (Air Sea Containers) to
540 the Kennedy Institute of Rheumatology by Royal Mail delivery. Samples were immediately stored
541 at -80°C upon arrival until processed for DNA and RNA extractions. For fresh-frozen stool used
542 for bacterial isolation, participants were provided with stool collection kits during clinic visits.
543 Samples were collected at home using a FecesCatcher (Tag Hemi) and a feces container
544 (Starstedt) and placed inside a disposable styrofoam container with frozen ice packs. Participants
545 returned the kits to the clinic in person and samples were frozen at -80°C immediately upon arrival.
546 Fresh-frozen stool was pulverized into a homogenous mixture utilizing a Biopulverizer (Stratech)
547 cooled in liquid nitrogen prior to aliquoting.

Statistical analysis
549 All raw, individual-level data for experiments where n<20 are presented in data file S2. Two
550 primary classes of statistical testing were used throughout this analysis, omnibus tests and per-

551   feature tests. The former assessed whether whole microbial community structure was significantly
552   different based on phenotype, whereas the latter assessed this for each individual feature (e.g.,
553   taxon, pathway, etc.). Omnibus tests were generally carried out using Bray-Curtis-based
554   PERMANOVA for the taxonomic, pathway and, enzyme-based feature tables using the vegan
555   v2.5-6 package(*107*) in R. With one expectation, we ran all models in an adjusted univariate
556   format, consisting of sequencing batch followed by the variable of interest [adonis(bray ≈ batch +
557   x)] with 1,000 permutations. However, to test the collection center variable, we added diagnosis
558   to the adjustments to the model, since diagnosis was substantially confounded with clinical site
559   (patients carrying certain diagnoses were only seen at certain centers). All p-values that are
560   presented are FDR corrected using the *p.adjust* function unless otherwise stated.

561   Additionally, for diagnosis and inflammation status, we used the package default pairwise.adonis
562   v0.0.1(*108*) to identify which of the diagnoses or inflammation status categories might be driving
563   the overall results. The same method PERMANOVA model was used for strain testing, but with
564   Kimura 2-parameter distances as input (ape::dist.dna). For our PanPhlAn presence/absence
565   data, we tested for significantly different presence or absence of genes using a chi-squared test
566   by leveraging the Gtest function in the DescTools package(*109*) in R. Difference in means for
567   both Kimura 2-parameter distances and changes in disease score and joint diameter were
568   quantified using Wilcoxon tests.

569   For parametric feature-wise multivariable testing we used MaAsLin 2 v1.4.0(*110*) in R, which finds
570   associations between microbial features and metadata of interest. MaAsLin uses a transformed
571   generalized linear model to associate each feature iteratively with covariates of interest, here
572   using a variance-stabilizing log transformation plus a small pseudocount of half the minimum
573   feature value for microbial relative abundances (total sum scaling). It then models each microbial
574   feature as a function of the patient's age and adjusts the resulting p-values for multiple hypothesis
575   tests, using a Benjamini–Hochberg correction. As noted above for different analyses, we used
576   several variants of the main feature-covariate model, where Feature ≈ batch + drug + age +
577   inflammation status OR patient diagnosis OR anemia status: first, a fully multivariable model that
578   was the most conservative adjusting for the most patient information; second, a reduced model
579   accounting only for technical sequencing batch; and finally, the least conservative model only
580   adjusting for age with the metadata of interest (inflammation status or patient diagnosis or anemia
581   status). In general, the model without sequencing batch was not used, since there was little
582   evidence of technical batch effects and since inflammation and diagnosis were imbalanced across
583   our sequencing batches. Further we compared the categorized way of processing CRP data to
584   other disease specific markers including BASDAI and the DAS28 metrics and they identified many
585   of the same taxonomic features and importantly in the same direction (**fig. S6**).

586   Most visualizations were carried out using standard methods in R ggplot2 v3.3.2(*111*), ggridges
587   v0.5.2(*112*), ggthemes v4.2.0(*111*), gridExtra v2.3(*113*), gtools v3.81(*114*), and ggtree
588   v2.0.2(*115, 116*). The principal coordinate analysis was done using the capscale function in
589   vegan(*107*). Additionally, we used GraPhlAn v0.9.7(*117*) to construct the cladogram in **Fig. 2**.
590   Additional R packages used for data manipulation and processing include pylr v1.8.6(*118*), dpylr
591   v1.0.2(*119*), scales v1.1.0(*119*), mgsub v1.71(*120*), and RColorBrewer v1.1.2(*121*).

592   # Supplemental Material

593   Fig. S1 to S26
594   Table S1 and S2
595   MDAR Reproducibility Checklist
596   Data file S1 and S2
597   References (122-146)

References

599 1. J. Lloyd-Price *et al.*, Multi-omics of the gut microbial ecosystem in inflammatory bowel
600 diseases. *Nature* **569**, 655-662 (2019).
601 2. G. Nakatsu *et al.*, Gut mucosal microbiome across stages of colorectal carcinogenesis.
602 *Nat Commun* **6**, 8727 (2015).
603 3. A. D. Kostic *et al.*, The dynamics of the human infant gut microbiome in development and
604 in progression toward type 1 diabetes. *Cell Host Microbe* **17**, 260-273 (2015).
605 4. J. Henao-Mejia *et al.*, Inflammasome-mediated dysbiosis regulates progression of NAFLD
606 and obesity. *Nature* **482**, 179-185 (2012).
607 5. S. Safiri *et al.*, Global, regional and national burden of rheumatoid arthritis 1990-2017: a
608 systematic analysis of the Global Burden of Disease study 2017. *Ann Rheum Dis* **78**,
609 1463-1471 (2019).
610 6. J. T. Liu, H. M. Yeh, S. Y. Liu, K. T. Chen, Psoriatic arthritis: Epidemiology, diagnosis, and
611 treatment. *World J Orthop* **5**, 537-543 (2014).
612 7. L. E. Dean *et al.*, Global prevalence of ankylosing spondylitis. *Rheumatology (Oxford)* **53**,
613 650-657 (2014).
614 8. H. J. Wu *et al.*, Gut-residing segmented filamentous bacteria drive autoimmune arthritis
615 via T helper 17 cells. *Immunity* **32**, 815-827 (2010).
616 9. G. T. Nepom *et al.*, HLA genes associated with rheumatoid arthritis. Identification of
617 susceptibility alleles using specific oligonucleotide probes. *Arthritis Rheum* **32**, 15-21
618 (1989).
619 10. B. Mesko *et al.*, Peripheral blood gene expression patterns discriminate among chronic
620 inflammatory diseases and healthy controls and identify novel targets. *BMC Med*
621 *Genomics*, 3-15 (2010).
622 11. A. H. M. Helm-van Mil, J. Z. Wesoly, T. W. Huizinga, Understanding the genetic
623 contribution to rheumatoid arthritis. *Curr Opin Rheumatol* **17**, 299-304 (2005).
624 12. J. Kurko *et al.*, Genetics of rheumatoid arthritis - a comprehensive review. *Clin Rev Allergy*
625 *Immunol* **45**, 170-179 (2013).
626 13. A. J. MacGregor *et al.*, Characterizing the quantitative genetic contribution to rheumatoid
627 arthritis using data from twins. *Arthritis Rheum* **43**, 30-37 (2000).
628 14. D. A. Brewerton *et al.*, Ankylosing spondylitis and HL-A 27. *Lancet* **1**, 904-907 (1973).
629 15. U. Kiltz *et al.*, Do patients with non-radiographic axial spondylarthritis differ from patients
630 with ankylosing spondylitis? *Arthritis Care Res (Hoboken)* **64**, 1415-1422 (2012).
631 16. M. Rudwaleit *et al.*, The early disease stage in axial spondylarthritis: results from the
632 German Spondyloarthritis Inception Cohort. *Arthritis Rheum* **60**, 717-727 (2009).
633 17. L. Klareskog, L. Padyukov, L. Alfredsson, Smoking as a trigger for inflammatory rheumatic
634 diseases. *Curr Opin Rheumatol* **19**, 49-54 (2007).
635 18. D. D. Chaplin, Overview of the immune response. *J Allergy Clin Immunol* **125**, S3-23
636 (2010).
637 19. G. A. Bannatyne, A. S. Wohlmann, RHEUMATOID ARTHRITIS : ITS CLINICAL
638 HISTORY, ETIOLOGY, AND PATHOLOGY. *Lancet* **147**, 1120-1125 (1896).
639 20. J. D. Taurog *et al.*, The germfree state prevents development of gut and joint inflammatory
640 disease in HLA-B27 transgenic rats. *J Exp Med* **180**, 2359-2364 (1994).
641 21. H. C. Rath *et al.*, Normal luminal bacteria, especially Bacteroides species, mediate chronic
642 colitis, gastritis, and arthritis in HLA-B27/human beta2 microglobulin transgenic rats. *J Clin*
643 *Invest* **98**, 945-953 (1996).
644 22. Z. Sinkorova, J. Capkova, J. Niederlova, R. Stepankova, J. Sinkora, Commensal intestinal
645 bacterial strains trigger ankylosing enthesopathy of the ankle in inbred B10.BR (H-2(k))
646 male mice. *Hum Immunol* **69**, 845-850 (2008).

647 23. S. Abdollahi-Roodsaz *et al.*, Stimulation of TLR2 and TLR4 differentially skews the
648 balance of T cells in a mouse model of arthritis. *J Clin Invest* **118**, 205-216 (2008).

649 24. M. F. Van den Broek, L. B. Van de Putte, W. B. Van den Berg, Crohn's disease associated
650 with arthritis: a possible role for cross-reactivity between gut bacteria and cartilage in the
651 pathogenesis of arthritis. *Arthritis Rheum* **31**, 1077-1079 (1988).

652 25. P. E. Pease, H. Chahal, J. E. Tallack, M. R. Lane, R. N. Allan, Cross-reactivity studies on
653 bacteria believed to be associated with inflammatory bowel disease (IBD), ankylosing
654 spondylitis (AS) and reactive arthritis (ReA). *Br J Rheumatol* **27 Suppl 2**, 32-33 (1988).

655 26. S. N. Lichtman *et al.*, Reactivation of arthritis induced by small bowel bacterial overgrowth
656 in rats: role of cytokines, bacteria, and bacterial polymers. *Infect Immun* **63**, 2295-2301
657 (1995).

658 27. E. Eerola *et al.*, Intestinal flora in early rheumatoid arthritis. *Br J Rheumatol* **33**, 1030-1038
659 (1994).

660 28. J. Vaahtovuo, E. Munukka, M. Korkeamaki, R. Luukkainen, P. Toivanen, Fecal microbiota
661 in early rheumatoid arthritis. *J Rheumatol* **35**, 1500-1505 (2008).

662 29. J. U. Scher *et al.*, Expansion of intestinal Prevotella copri correlates with enhanced
663 susceptibility to arthritis. *Elife* **2**, e01202 (2013).

664 30. S. Stebbings *et al.*, Comparison of the faecal microflora of patients with ankylosing
665 spondylitis and controls using molecular methods of analysis. *Rheumatology (Oxford)* **41**,
666 1395-1401 (2002).

667 31. M. E. Costello *et al.*, Brief Report: Intestinal Dysbiosis in Ankylosing Spondylitis. *Arthritis
668 Rheumatol* **67**, 686-691 (2015).

669 32. M. L. Stoll *et al.*, Altered microbiota associated with abnormal humoral immune responses
670 to commensal organisms in enthesitis-related arthritis. *Arthritis Res Ther* **16**, 486 (2014).

671 33. J. U. Scher *et al.*, Decreased bacterial diversity characterizes the altered gut microbiota in
672 patients with psoriatic arthritis, resembling dysbiosis in inflammatory bowel disease.
673 *Arthritis Rheumatol* **67**, 128-139 (2015).

674 34. X. Zhang *et al.*, The oral and gut microbiomes are perturbed in rheumatoid arthritis and
675 partly normalized after treatment. *Nat Med* **21**, 895-905 (2015).

676 35. C. Wen *et al.*, Quantitative metagenomics reveals unique gut microbiome biomarkers in
677 ankylosing spondylitis. *Genome Biol* **18**, 142 (2017).

678 36. R. Huang *et al.*, Metagenome-wide association study of the alterations in the intestinal
679 microbiome composition of ankylosing spondylitis patients and the effect of traditional and
680 herbal treatment. *J Med Microbiol* **69**, 797-805 (2020).

681 37. J. Yin *et al.*, Shotgun metagenomics reveals an enrichment of potentially cross-reactive
682 bacterial epitopes in ankylosing spondylitis patients, as well as the effects of TNFi therapy
683 upon microbiome composition. *Ann Rheum Dis* **79**, 132-140 (2020).

684 38. C. Zhou *et al.*, Metagenomic profiling of the pro-inflammatory gut microbiota in ankylosing
685 spondylitis. *J Autoimmun* **107**, 102360 (2020).

686 39. T. Kishikawa *et al.*, Metagenome-wide association study of gut microbiome revealed novel
687 aetiology of rheumatoid arthritis in the Japanese population. *Ann Rheum Dis* **79**, 103-111
688 (2020).

689 40. M. Berland *et al.*, Both Disease Activity and HLA-B27 Status Are Associated With Gut
690 Microbiome Dysbiosis in Spondyloarthritis Patients. *Arthritis Rheumatol*, (2022).

691 41. V. K. Gupta *et al.*, Gut microbial determinants of clinically important improvement in
692 patients with rheumatoid arthritis. *Genome Med* **13**, 149 (2021).

693 42. J. U. Scher, D. R. Littman, S. B. Abramson, Microbiome in Inflammatory Arthritis and
694 Human Rheumatic Diseases. *Arthritis Rheumatol* **68**, 35-45 (2016).

695 43. Y. Zhao *et al.*, Hidden link in gut-joint axis: gut microbes promote rheumatoid arthritis at
696 early stage by enhancing ascorbate degradation. *Gut* **71**, 1041-1043 (2022).

697 44. H. W. van Steenbergen *et al.*, EULAR definition of arthralgia suspicious for progression to
698   rheumatoid arthritis. *Ann Rheum Dis* **76**, 491-496 (2017).
699 45. M. Asquith *et al.*, HLA Alleles Associated With Risk of Ankylosing Spondylitis and
700   Rheumatoid Arthritis Influence the Gut Microbiome. *Arthritis Rheumatol* **71**, 1642-1650
701   (2019).
702 46. H. Xu, J. Yin, HLA risk alleles and gut microbiome in ankylosing spondylitis and
703   rheumatoid arthritis. *Best Pract Res Clin Rheumatol* **33**, 101499 (2019).
704 47. M. M. Newkirk, A. Zbar, M. Baron, A. R. Manges, Distinct bacterial colonization patterns
705   of Escherichia coli subtypes associate with rheumatoid factor status in early inflammatory
706   arthritis. *Rheumatology (Oxford)* **49**, 1311-1316 (2010).
707 48. A. Pianta *et al.*, Evidence of the Immune Relevance of Prevotella copri, a Gut Microbe, in
708   Patients With Rheumatoid Arthritis. *Arthritis Rheumatol* **69**, 964-975 (2017).
709 49. A. Kojima *et al.*, Infection of specific strains of Streptococcus mutans, oral bacteria,
710   confers a risk of ulcerative colitis. *Sci Rep* **2**, 332 (2012).
711 50. G. Kaci *et al.*, Anti-inflammatory properties of Streptococcus salivarius, a commensal
712   bacterium of the oral cavity and digestive tract. *Appl Environ Microbiol* **80**, 928-934 (2014).
713 51. J. U. Scher *et al.*, Periodontal disease and the oral microbiota in new-onset rheumatoid
714   arthritis. *Arthritis Rheum* **64**, 3083-3094 (2012).
715 52. L. Maier *et al.*, Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature*
716   **555**, 623-628 (2018).
717 53. F. Imhann *et al.*, Proton pump inhibitors affect the gut microbiome. *Gut* **65**, 740-748 (2016).
718 54. P. de Pablo, T. Dietrich, T. E. McAlindon, Association of periodontal disease and tooth
719   loss with rheumatoid arthritis in the US population. *J Rheumatol* **35**, 70-76 (2008).
720 55. R. Martin *et al.*, Faecalibacterium prausnitzii prevents physiological damages in a chronic
721   low-grade inflammation murine model. *BMC Microbiol* **15**, 67 (2015).
722 56. C. T. Brown *et al.*, Gut microbiome metagenomics analysis suggests a functional model
723   for the development of autoimmunity for type 1 diabetes. *PLoS One* **6**, e25792 (2011).
724 57. K. Machiels *et al.*, A decrease of the butyrate-producing species Roseburia hominis and
725   Faecalibacterium prausnitzii defines dysbiosis in patients with ulcerative colitis. *Gut* **63**,
726   1275-1283 (2014).
727 58. B. J. Parker, P. A. Wearsch, A. C. M. Veloo, A. Rodriguez-Palacios, The Genus Alistipes:
728   Gut Bacteria With Emerging Implications to Inflammation, Cancer, and Mental Health.
729   *Front Immunol* **11**, 906 (2020).
730 59. P. Louis, P. Young, G. Holtrop, H. J. Flint, Diversity of human colonic butyrate-producing
731   bacteria revealed by analysis of the butyryl-CoA:acetate CoA-transferase gene. *Environ*
732   *Microbiol* **12**, 304-314 (2010).
733 60. S. L. La Rosa *et al.*, The human gut Firmicute Roseburia intestinalis is a primary degrader
734   of dietary beta-mannans. *Nat Commun* **10**, 905 (2019).
735 61. A. Picchianti-Diamanti *et al.*, Analysis of Gut Microbiota in Rheumatoid Arthritis Patients:
736   Disease-Related Dysbiosis and Modifications Induced by Etanercept. *Int J Mol Sci* **19**,
737   (2018).
738 62. R. Bodkhe, B. Balakrishnan, V. Taneja, The role of microbiome in rheumatoid arthritis
739   treatment. *Ther Adv Musculoskelet Dis* **11**, 1759720x19844632 (2019).
740 63. J. Chen *et al.*, An expansion of rare lineage intestinal microbes characterizes rheumatoid
741   arthritis. *Genome Med* **8**, 43 (2016).
742 64. M. Zhang *et al.*, Faecalibacterium prausnitzii inhibits interleukin-17 to ameliorate colorectal
743   colitis in rats. *PLoS One* **9**, e109146 (2014).
744 65. H. M. Hamer *et al.*, Review article: the role of butyrate on colonic function. *Aliment*
745   *Pharmacol Ther.* **27**, 104-119 (2008).

66. S. Yuille, N. Reichardt, S. Panda, H. Dunbar, I. E. Mulder, Human gut bacteria as potent class I histone deacetylase inhibitors in vitro through production of butyric acid and valeric acid. *PLoS One* **13**, e0201073 (2018).

67. X. C. Morgan *et al.*, Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* **13**, R79 (2012).

68. K. Atarashi *et al.*, Ectopic colonization of oral bacteria in the intestine drives T(H)1 cell induction and inflammation. *Science* **358**, 359-365 (2017).

69. C. W. Png *et al.*, Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. *Am J Gastroenterol* **105**, 2420-2428 (2010).

70. E. H. Crost *et al.*, The mucin-degradation strategy of Ruminococcus gnavus: The importance of intramolecular trans-sialidases. *Gut Microbes* **7**, 302-312 (2016).

71. D. Azzouz *et al.*, Lupus nephritis is linked to disease-activity associated expansions and immunity to a gut commensal. *Ann Rheum Dis* **78**, 947-956 (2019).

72. H. H. Chua *et al.*, Intestinal Dysbiosis Featuring Abundance of Ruminococcus gnavus Associates With Allergic Diseases in Infants. *Gastroenterology* **154**, 154-167 (2018).

73. E. Coletto *et al.*, The role of the mucin-glycan foraging Ruminococcus gnavus in the communication between the gut and the brain. *Gut Microbes* **14**, 2073784 (2022).

74. M. Breban *et al.*, Faecal microbiota study reveals specific dysbiosis in spondyloarthritis. *Ann Rheum Dis* **76**, 1614-1622 (2017).

75. A. B. Hall *et al.*, A novel Ruminococcus gnavus clade enriched in inflammatory bowel disease patients. *Genome Med* **9**, 103 (2017).

76. M. T. Henke *et al.*, Ruminococcus gnavus, a member of the human gut microbiome associated with Crohn's disease, produces an inflammatory polysaccharide. *Proc Natl Acad Sci U S A* **116**, 12672-12677 (2019).

77. M. Schirmer, A. Garner, H. Vlamakis, R. J. Xavier, Microbial genes and pathways in inflammatory bowel disease. *Nat Rev Microbiol* **17**, 497-511 (2019).

78. D. T. Truong, A. Tett, E. Pasolli, C. Huttenhower, N. Segata, Microbial strain-level population structure and genetic diversity from metagenomes. *Genome Res* **27**, 626-638 (2017).

79. M. Scholz *et al.*, Strain-level microbial epidemiology and population genomics from shotgun metagenomics. *Nat Methods* **13**, 435-438 (2016).

80. M. T. Henke *et al.*, Capsular polysaccharide correlates with immune response to the human gut microbe Ruminococcus gnavus. *Proc Natl Acad Sci U S A* **118**, (2021).

81. D. O. Kennedy, B Vitamins and the Brain: Mechanisms, Dose and Efficacy--A Review. *Nutrients* **8**, 68 (2016).

82. Q. Wang, R. Xu, Data-driven multiple-level analysis of gut-microbiome-immune-joint interactions in rheumatoid arthritis. *BMC Genomics* **20**, 124 (2019).

83. P. Spanel, D. Smith, Volatile compounds in health and disease. *Curr Opin Clin Nutr Metab Care* **14**, 455-460 (2011).

84. R. M. Thorn, J. Greenman, Microbial volatile compounds in health and disease conditions. *J Breath Res* **6**, 024001 (2012).

85. A. Marcuzzi *et al.*, Natural isoprenoids are able to reduce inflammation in a mouse model of mevalonate kinase deficiency. *Pediatr Res* **64**, 177-182 (2008).

86. D. Yu *et al.*, The Gut Microbiome and Metabolites Are Altered and Interrelated in Patients With Rheumatoid Arthritis. *Front Cell Infect Microbiol* **11**, 763507 (2021).

87. R. Sinha *et al.*, Assessment of variation in microbial community amplicon sequencing by the Microbiome Quality Control (MBQC) project consortium. *Nat Biotechnol* **35**, 1077-1086 (2017).

88. S. Ma *et al.*, Population structure discovery in meta-analyzed microbial communities and inflammatory bowel disease using MMUPHin. *Genome Biol* **23**, 208 (2022).

796   89.   M. Blits *et al.*, Methotrexate normalizes up-regulated folate pathway genes in rheumatoid
797   arthritis. *Arthritis Rheum* **65**, 2791-2802 (2013).
798   90.   A. L. Bayer, C. A. Fraker, The Folate Cycle As a Cause of Natural Killer Cell Dysfunction
799   and Viral Etiology in Type 1 Diabetes. *Front Endocrinol (Lausanne)* **8**, 315 (2017).
800   91.   D. N. Meadows, R. H. Bahous, A. F. Best, R. Rozen, High Dietary Folate in Mice Alters
801   Immune Response and Reduces Survival after Malarial Infection. *PLoS One* **10**,
802   e0143738 (2015).
803   92.   C. Courtemanche, I. Elson-Schwab, S. T. Mashiyama, N. Kerry, B. N. Ames, Folate
804   deficiency inhibits the proliferation of primary human CD8+ T lymphocytes in vitro. *J
805   Immunol* **173**, 3186-3192 (2004).
806   93.   M. Kinoshita *et al.*, Dietary folic acid promotes survival of Foxp3+ regulatory T cells in the
807   colon. *J Immunol* **189**, 2869-2878 (2012).
808   94.   T. Ganz, E. Nemeth, Iron sequestration and anemia of inflammation. *Semin Hematol* **46**,
809   387-393 (2009).
810   95.   N. K. Das *et al.*, Microbial Metabolite Signaling Is Required for Systemic Iron Homeostasis.
811   *Cell Metab* **31**, 115-130.e116 (2020).
812   96.   P. L. Carver, The Battle for Iron between Humans and Microbes. *Curr Med Chem* **25**, 85-
813   96 (2018).
814   97.   J. P. Motta *et al.*, Iron Sequestration in Microbiota Biofilms As A Novel Strategy for Treating
815   Inflammatory Bowel Disease. *Inflamm Bowel Dis* **24**, 1493-1502 (2018).
816   98.   A. Wilson, H. T. Yu, L. T. Goodnough, A. R. Nissenson, Prevalence and outcomes of
817   anemia in rheumatoid arthritis: a systematic review of the literature. *Am J Med* **116 Suppl
818   7A**, 50s-57s (2004).
819   99.   L. Niccoli, C. Nannini, E. Cassara, O. Kaloudi, F. Cantini, Frequency of anemia of
820   inflammation in patients with ankylosing spondylitis requiring anti-TNFalpha drugs and
821   therapy-induced changes. *Int J Rheum Dis* **15**, 56-61 (2012).
822   100.   T. K. Kvien, T. Uhlig, S. Odegard, M. S. Heiberg, Epidemiological aspects of rheumatoid
823   arthritis: the sex ratio. *Ann N Y Acad Sci* **1069**, 212-222 (2006).
824   101.   A. T. Masi, W. R. Wilkins, Does male:female sex ratio in ankylosing spondylitis change
825   with age? *J Rheumatol* **23**, 947-948 (1996).
826   102.   A framework for human microbiome research. *Nature* **486**, 215-221 (2012).
827   103.   P. A. Harris *et al.*, Research electronic data capture (REDCap)--a metadata-driven
828   methodology and workflow process for providing translational research informatics
829   support. *J Biomed Inform* **42**, 377-381 (2009).
830   104.   P. A. Harris *et al.*, The REDCap consortium: Building an international community of
831   software platform partners. *J Biomed Inform* **95**, 103208 (2019).
832   105.   M. Rudwaleit *et al.*, The Assessment of SpondyloArthritis International Society
833   classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general.
834   *Ann Rheum Dis* **70**, 25-31 (2011).
835   106.   E. A. Franzosa *et al.*, Relating the metatranscriptome and metagenome of the human gut.
836   *Proc Natl Acad Sci U S A* **111**, E2329-2338 (2014).
837   107.   P. Dixon, VEGAN, a package of R functions for community ecology. *Journal of Vegetation
838   Science*, (2003).
839   108.   P. M. Arbizu, pairwiseAdonis: Pairwise Multilevel Comparison using Adonis. . *R package
840   version 0.0.1.*, (2017).

841   109.   A. Signorell, e. m. al., DescTools: Tools for descriptive statistics. *R package version
842   0.99.34.*, (2020).
843   110.   H. Mallick *et al.*, Multivariable Association Discovery in Population-scale Meta-omics
844   Studies. *PLoS Comput Biol* **17**, e1009442 (2021).

845 111. J. B. Arnold, ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'. *R package version*
846 *4.2.0*, (2019).
847 112. C. O. Wilke, ggridges: Ridgeline Plots in 'ggplot2'. *R package version 0.5.2*, (2020).
848 113. B. Auguie, gridExtra: Miscellaneous Functions for "Grid" Graphics. *R package version*
849 *2.3.*, (2017).
850 114. G. R. Warnes, B. Bolker, T. Lurnley, gtools: Various R Programming Tools. *R package*
851 *version 3.8.1.*, (2018).
852 115. G. Yu, D. K. Smith, H. Zhu, Y. Guan, T. T. Lam, ggtree: an r package for visualization and
853 annotation of phylogenetic trees with their covariates and other associated data. *Methods*
854 *in Ecology and Evolution* **8**, 28-36 (2016).
855 116. G. Yu, T. T. Lam, H. Zhu, Y. Guan, Two Methods for Mapping and Visualizing Associated
856 Data on Phylogeny Using Ggtree. *Mol Biol Evol* **35**, 3041-3043 (2018).
857 117. F. Asnicar, G. Weingart, T. L. Tickle, C. Huttenhower, N. Segata, Compact graphical
858 representation of phylogenetic data and metadata with GraPhlAn. *PeerJ* **3**, (2015).
859 118. H. Wickham, The Split-Apply-Combine Strategy for Data Analysis. *Journal of Statistical*
860 *Software* **40**, (2011).
861 119. H. Wickham, R. Francois, L. Henry, K. Muller, dplyr: A Grammar of Data Manipulation. .
862 *R package version 1.0.2*, (2020).

863 120. M. Ewing, mgsub: Safe, Multiple, Simultaneous String Substitution. . *R package version*
864 *1.7.1*, (2019).

865 121. Neuwirth, RColorBrewer: ColorBrewer Palettes. *R package version 1.1-2.*, (2014).

866 ## Acknowledgements

884 **Author contributions**

885 This study was conceptualized and designed by FP, CH, KR, PB, AGP, DRL, SPY, JJF, LRW
886 and PCT. Data and specimen collection was completed by PC, SJB, AF, CP, SK, AGP, PB, and
887 KR. Molecular methods prior to sequencing were done by PC, IB, CP, CM, VK, and LHL.

888 Computational storage of clinical and sequence data was designed, maintained, and organized
889 by KNT, KSB, NEI, PC, LJM, MAJ, and BM. Analysis was completed by KNT, KSB, NEI, GJB,
890 SM, and LHN. Feedback and directions were guided by AGP, PB, and KR. Mouse work was
891 completed by GJB and JJF. Manuscript was written by KNT, LHN, AP and CH. Editing was
892 completed by all authors.

Competing interests

Data availability

All data associated with this study are in the paper or supplementary materials. Sequence data
and metadata are available for approved user to download through the EGA (https://ega-archive.org/), study accession number EGAS00001005525. Processed taxonomic and functional
tables are available in data file S1. Bioinformatic workflows for metagenomic processing are
available at https://huttenhower.sph.harvard.edu/biobakery_workflows, these include some basic
statistical and visualization scripts. Custom analysis scripts are available at
http://huttenhower.sph.harvard.edu/Adult_cross-sectional_IAMC

The Inflammatory Arthritis Microbiome Consortium (IAMC) investigators group

907 In addition to IAMC members who are authors (Peter C Taylor[9], Lucy R Wedderburn[10,11,12],
908 Stephen P Young[13], Dan R Littman[14], Jeremiah J Faith[5], Paul Bowness[9], Karim Raza[7,8,13], Fiona
909 Powrie[4], and Curtis Huttenhower[1,2,3,18]), the following IAMC members are collaborators who have
910 contributed to study design, data analysis, and interpretation:

911 Coziana Ciurtin[12,19] and Claudia Mauri[20]

913 Affiliations 1 to 18 can be found on the first page of the paper.

914 [19]Division of Medicine, University College London, London, WC1E 6BT, UK

915 [20]Centre for Rheumatology and Bloomsbury Rheumatology Unit, UCL School of Life and Medical
916 Science, London, WC1E 6BT, UK