






OPEN ACCESS

TRANSLATIONAL SCIENCE

Stage-specific roles of microbial dysbiosis and metabolic disorders in rheumatoid arthritis

Mingyue Cheng ^{1,2}, Yan Zhao,¹ Yazhou Cui,¹ Chaofang Zhong,² Yuguo Zha,² Shufeng Li,¹ Guangxiang Cao,¹ Mian Li,¹ Lei Zhang,³ Kang Ning ^{1,2}, Jinxiang Han ¹

Handling editor Josef S Smolen

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/ard-2022-222871>).

For numbered affiliations see end of article.

Correspondence to

Dr Jinxiang Han, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan, Shandong, China; jxhan@sdfmu.edu.cn, Dr Kang Ning, Huazhong University of Science and Technology, Wuhan, Hubei, China; ningkang@hust.edu.cn and Dr Lei Zhang, Shandong University, Jinan, Shandong, China; zhanglei7@sdu.edu.cn

MC and YZ contributed equally.

MC and YZ are joint first authors.

LZ, KN and JH are joint senior authors.

Received 31 May 2022

Accepted 7 August 2022

Published Online First

19 August 2022

ABSTRACT

Objective Rheumatoid arthritis (RA) is a progressive disease including four stages, where gut microbiome is associated with pathogenesis. We aimed to investigate stage-specific roles of microbial dysbiosis and metabolic disorders in RA.

Methods We investigated stage-based profiles of faecal metagenome and plasma metabolome of 76 individuals with RA grouped into four stages (stages I–IV) according to 2010 RA classification criteria, 19 individuals with osteoarthritis and 27 healthy individuals. To verify bacterial invasion of joint synovial fluid, 16S rRNA gene sequencing, bacterial isolation and scanning electron microscopy were conducted on another validation cohort of 271 patients from four RA stages.

Results First, depletion of *Bacteroides uniformis* and *Bacteroides plebeius* weakened glycosaminoglycan metabolism ($p < 0.001$), continuously hurting articular cartilage across four stages. Second, elevation of *Escherichia coli* enhanced arginine succinyltransferase pathway in the stage II and stage III ($p < 0.001$), which was correlated with the increase of the rheumatoid factor ($p = 1.35 \times 10^{-3}$) and could induce bone loss. Third, abnormally high levels of methoxyacetic acid ($p = 1.28 \times 10^{-8}$) and cysteine-S-sulfate ($p = 4.66 \times 10^{-12}$) inhibited osteoblasts in the stage II and enhanced osteoclasts in the stage III, respectively, promoting bone erosion. Fourth, continuous increase of gut permeability may induce gut microbial invasion of the joint synovial fluid in the stage IV.

Conclusions Clinical microbial intervention should consider the RA stage, where microbial dysbiosis and metabolic disorders present distinct patterns and played stage-specific roles. Our work provides a new insight in understanding gut–joint axis from a perspective of stages, which opens up new avenues for RA prognosis and therapy.

INTRODUCTION

Rheumatoid arthritis (RA) affects over tens of millions of people worldwide.¹ RA is recognised clinically as a progressive, inflammatory and auto-immune disease that primarily affects the joints and typically has four stages^{2–5}: (1) In the first stage, the synovium of the joints is inflamed and most people have minor symptoms such as stiffness on awakening; (2) In the second stage, the inflamed synovium has caused damage to the joint cartilage and people begin to feel swelling, and have a

WHAT IS ALREADY KNOWN ON THIS TOPIC

- ⇒ Rheumatoid arthritis (RA) is a progressive disease, clinically including four stages.
- ⇒ Intestinal microbiome is associated with the pathogenesis of RA.
- ⇒ Joint synovial fluid is generally considered as sterile.

WHAT THIS STUDY ADDS

- ⇒ This is the first study focusing on RA stages to report microbial and metabolic profiles and roles, particularly their enhancement of inflammation, bone loss and bone erosion in the stage II and stage III.
- ⇒ Joint synovial fluid is not sterile, where bacterial invasion happened in the stage IV.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- ⇒ The study provides microbial and metabolic targets for each stage of RA.
- ⇒ Further experiments and intervention on microbiota of joint synovial fluid are warranted for patients in the stage IV.

restricted range of motion; (3) In the third stage, RA has proceeded to a severe state when bone erosion begins and the cartilage on the surface of the bones has deteriorated, resulting in the bones rubbing against one another and (4) In the fourth stage, certain joints are severely deformed and lose function. To inhibit RA progression, specific therapeutic strategies are necessary for people across different RA stages.

Gut microbial dysbiosis has been implicated in the pathogenesis of RA via a range of mechanisms such as metabolic perturbation and immune response regulation, which is known as the gut–joint axis,^{6,7} for instance, increased abundance of *Prevotella* and *Collinsella* in patients with RA are correlated with the production of T_H17 cell cytokines.^{8,9} Moreover, Gut microbes and their products were likely to be transited to the joint due to the increased gut permeability.⁶ Metabolites have also been correlated to immunity regulation in RA: administration of Short-chain fatty acids to mice with collagen-induced arthritis (CIA) can reduce the severity of arthritis by modulation of IL-10.^{6,10} Comprehensive metagenomic and metabolomic analyses could



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Cheng M, Zhao Y, Cui Y, et al. *Ann Rheum Dis* 2022;**81**:1669–1677.

Table 1 General characteristics at stool collection of multiomics cohort (122 participants)

	Patients with RA from four stages				OA n=19	HC n=27
	RAS1 n=15	RAS2 n=21	RAS3 n=18	RAS4 n=22		
Age (years), median (IQR)	52 (50–60)	64 (59–67)	59 (50–66)	60 (54–66)	66 (64–71)	56 (50–60)
Female sex, n (%)	12 (80)	16 (76)	12 (67)	21 (95)	15 (79)	19 (70)
Classification score						
A (IQR)	3 (3–3)	3 (3–3)	5 (5–5)	5 (5–5)	1 (0–1)	0
B (IQR)	2 (2–2)	3 (3–3)	3 (3–3)	3 (3–3)	0	0
C (IQR)	1 (1–1)	1 (1–1)	1 (1–1)	1 (1–1)	1 (1–1)	0
D (IQR)	1 (1–1)	1 (1–1)	0 (0–0)	1 (1–1)	1 (1–1)	0
Sum score (IQR)	7 (7–7)	8 (8–8)	9 (9–9)	10 (10–10)	3 (2–3)	0
ACPA positivity, n (%)	5 (33)	18 (86)	13 (72)	17 (77)	0	
ESR (IQR)	60.00 (30.50–87.00)	72.00 (45.00–116.00)	48.00 (29.50–60.00)	60.00 (40.50–80.00)	23.00 (15.75–41.50)	
CRP (IQR)	29.10 (9.20–53.40)	43.00 (15.95–56.55)	38.60 (12.50–44.40)	27.15 (20.32–64.45)	17.45 (4.62–65.75)	
RF (IQR)	52.00 (39.50–60.50)	421.00 (186.00–590.00)	34.00 (19.50–80.00)	279.50 (106.00–359.00)	26.00 (22.00–29.00)	

ACPA positivity was defined as a concentration of greater than 5 µL/mL.

Classification scores were summarised according to 2010 RA classification criteria: A, joint involvement; B, serology; C, acute-phase reactants; D, duration of symptoms.

ACPA, anticitrullinated protein antibody; CRP, C reactive protein; ESR, erythrocyte sedimentation rate; HC, healthy controls; OA, osteoarthritis; RA, rheumatoid arthritis; RAS1–4, the first to fourth stage of RA; RF, rheumatoid factor.

therefore enhance our understanding about the gut–joint axis. However, the role of the gut–joint axis across successive stages of RA is understudied,^{6,11} where more examinations may provide an alternative approach to ameliorate RA progression.

Here, we aimed to investigate the stage-based profiles and roles of the gut–joint axis in RA pathogenesis, and whether or in which stage gut microbial invasion of the joint synovial fluid happened.

MATERIALS AND METHODS

Study design and sample collection

Data collection for this multiomics study was conducted in The First Affiliated Hospital of Shandong First Medical University (Jinan, Shandong, China), which was a provincial-level large-scale comprehensive tertiary first-class hospital and had tens of thousands of outpatients with arthritis per year. A total of 122 faecal and 122 plasma samples were collected from 122 outpatients of the The First Affiliated Hospital of Shandong First Medical University from 2017 to 2020.⁷ These outpatients included 76 patients with RA, 19 patients with OA and 27 healthy individuals (table 1). Patients with RA were grouped into four RA stages including RAS1 (n=15), RAS2 (n=21), RAS3 (n=18) and RAS4 (n=22) according to the rheumatoid diagnostic score,³ where RAS1, RAS2, RAS3 and RAS4 has a score of 6–7, 8, 9 and 10, respectively. The score was evaluated by the sum of four categories as summarised in the 2010 RA classification criteria.³ Faecal samples were collected and sequenced and plasma samples were used to test the plasma metabolites, anticitrullinated protein antibody, erythrocyte sedimentation rate, C reactive protein, rheumatoid factor, cytokines and plasma metabolites.

To confirm the bacterial invasion of the joint synovial fluid, another cohort of 271 with RA of four distinct stages were recruited, including 52 patients in RAS1, 66 in RAS2, 67 in RAS3 and 86 in RAS4. Synovial fluid samples were collected aseptically from knee joints during therapeutic aspiration. The entire experiment was conducted in a completely sterile atmosphere. For each patient, a total of 7 mL synovial fluid was collected, of which 5 mL was used for 16S rRNA gene sequencing, 1 mL was used for bacteria isolation and 1 mL synovial fluid was prepared for scanning electron microscopy.

All of the participants were at fasting status during the sample collection in the morning. Only participants who met the standard were recruited in this study: Recruited individuals had not received treatment in the recent month and were in the active period, and had no malignant tumour, no other rheumatic diseases such as ankylosing spondylitis, psoriasis, gout, no gastrointestinal diseases such as diarrhoea, constipation and haematochezia in the recent month, no infections, no other comorbidity such as diabetes and hepatitis B.

Metagenomic sequencing and processing to analyse the faecal microbiome

Whole-genome shotgun sequencing and processing of faecal samples, non-redundant gene catalogue construction, identification of metagenomic species (MGS), functional annotation to Kyoto Encyclopaedia of Genes and Genomes (KEGG) were performed (details in online supplemental text). Two parallel processes were used for gut metagenomic data analysis: One was based on 4 million non-redundant genes and investigated the functional composition across RA stages and OA, as well as the MGS that most drove the correlation of these microbial functions with RA or OA and (2) The other reported the 232 classified microbial species composition across RA stages and OA, profiled by MetaPhlan2¹² (V2.7.8).

UHPLC-QTOF-mass spectrometry analysis of plasma metabolites

Untargeted plasma metabolome was examined by ultra-performance liquid chromatography-quadrupole time-of-flight (UHPLC-QTOF) mass spectrometry: liquid chromatography with tandem mass spectrometry on an UHPLC system (1290, Agilent Technologies) with a UPLC BEH Amide column (1.7 µm 2.1×100 mm, Waters) coupled to TripleTOF 6600 (Q-TOF, AB Sciex) and QTOF 6550 (Agilent) (details in online supplemental text).

16S rRNA gene sequencing and processing to analyse the synovial fluid microbiota

Bacterial DNA was extracted from 271 5 mL synovial fluid samples. The tube containing PBS serves as environmental

control. Only a total of 86 synovial fluid samples from patients in RAS4 had enough bacteria DNA content (≥ 10 ng) (Bacterial DNA Kit, TIANGEN) for bacteria 16S rRNA gene high-throughput sequencing. The V1/V2 hypervariable regions of the 16S ribosomal RNA gene were sequenced using the Illumina HiSeq platform. The 16S sequence paired-end data set was joined and quality filtered using the FLASH as previously described.¹³ Taxonomic annotation was then performed (details in online supplemental text).

Bacterial isolation and scanning electron microscopy

Six synovial fluid samples (1 mL) per RA stage were used for bacteria isolation, and the obtained isolated colonies were identified using 16S rRNA gene sequencing (details in online supplemental text). For the samples from which bacterial can be isolated, synovial fluid samples (1 mL) of the same individuals were then filtered and imaged with scanning electron microscopy (ZEISS Sigma 300, details in online supplemental text).

Statistical analysis

Samples were divided into three groups including the healthy group, the OA group and the RA group. Samples of the RA group were further divided into four subgroups including RAS1, RAS2, RAS3, RAS4. For comparisons of vectors across groups or subgroups, such as microbial species abundance, KO abundance, metabolite intensity. Mann-Whitney-Wilcoxon test (p values) with Benjamini and Hochberg correction (q values) was used to test the significance. A threshold for statistical significance was $p < 0.05$, and for multiple testing the threshold was $p < 0.05$ and $q < 0.1$.

For correlations between KEGG modules and clinical phenotypes including arthritis (healthy=0, OA=1, RA=2), cytokine levels and rheumatic factor level, owing to that a KEGG module contained multiple KOs, Spearman correlation coefficients (SCC) between abundances of KOs and clinical phenotypes were first calculated. Subsequently, Mann-Whitney-Wilcoxon test (p values) with Benjamini and Hochberg correction (q values) was used to test if SCC between the KOs in a given KEGG module and phenotypes were different from that between all the other KOs out of the KEGG module and phenotypes. In this process, the KEGG module with statistical significance was viewed as significantly correlated with the clinical phenotypes. A threshold for statistical significance was $p < 0.05$ and $q < 0.1$. Considering that sex and age might have potential effects on gut microbiome,¹⁴ partial SCCs with age and gender adjusted were also calculated and compared, and a threshold for statistical significance was $p_{\text{partial}} < 0.05$ and $q_{\text{partial}} < 0.1$.

Leave-one-out analysis was used to test which MGS was driving the observed correlations between KEGG modules and arthritis. Owing to that one MGS contained multiple genes that were mapped to KOs, if one MGS was excluded in the dataset, the overall profiles of the KO abundance would change, resulting in the change of the correlations between KEGG modules and arthritis. Therefore, to determine the driving effects of each of MGS, the calculation of the KO abundance was iterated excluding the genes from a different MGS in each iteration, and the correlations between each KEGG module and arthritis were recalculated. Finally, the driving effects of a given MGS on a specified correlation was defined as the change in median SCC between KOs and arthritis when genes from the respective MGS were left out.

To determine the diagnostic potential of RA stages using multiomics features, random forest algorithm was performed on

6,224 KOs, 232 microbial species and 277 plasma metabolites, using the R package 'randomForest'. Function 'trainControl' in R package 'caret' was used to perform 10 repeats of 10-fold cross-validation for each data set. Function 'train' in R package 'caret' was used to fit models over different tuning parameters to determine the 'mtry' for random forest algorithm. Gini coefficients were used to measure how each variable contributed to the homogeneity of the nodes and leaves in the resulting random forest.

RESULTS

Stage-specific microbial taxonomic profiles

We obtained a total of 231 classified microbial species from metagenomic data, and tested their alterations in each stage of RA, as compared with healthy controls (see online supplemental figure S1, table S1–S5). The elevated species in RA progression were mostly from the phyla Firmicutes and Actinobacteria, while the depleted species were predominantly from the phylum Bacteroides ($q < 0.1$). We found certain microbes did not remain altered across RA stages, as compared with healthy controls. *Bifidobacterium dentium*, for instance, was reported to be associated with the development of dental caries and periodontal disease, both of which were particularly prevalent in patients with RA.^{15–16} Compared with healthy controls, it remained elevated across RA stages except for RAS1 (RAS2: $p = 7.16 \times 10^{-3}$, RAS3: $p = 3.70 \times 10^{-3}$, RAS4: $p = 9.15 \times 10^{-4}$). Moreover, we noticed that 29 species that were altered exclusively in a specific stage (see online supplemental table S1–S5). We found that *Collinsella aerofaciens* was elevated exclusively in RAS1 ($p = 0.043$). *C. aerofaciens* was previously reported to generate severe arthritis when inoculated into CIA-susceptible mice, and an in vitro experiment showed that *C. aerofaciens* could increase gut permeability and induce IL-17A expression, a key cytokine involved in RA pathogenesis.⁹ The elevation of *C. aerofaciens* in RAS1 might contribute to the early breach in gut barrier integrity, through which the translocation of microbial products would then trigger the subsequent clinical arthritis.⁶ Moreover, *Veillonella parvula*, whose infection could cause osteomyelitis,¹⁷ was found elevated exclusively in RAS3 ($p = 0.027$). *Eggerthella lenta* ($p = 0.018$) and *Bifidobacterium longum* ($p = 0.022$) were found elevated exclusively in RAS4. The gavage of *E. lenta* were reported to increase gut permeability and produce proinflammatory cytokines.¹⁸ We also recognised species altered exclusively in OA, such as elevated *Dialister invisus* ($p = 0.041$) that was positively correlated with spondyloarthritis severity.¹⁹ These stage-specific altered species had the potential to serve as the targets for intervention in a given RA stage.

Stage-specific microbial functional profiles

Next, we sought to detect the microbial dysfunction across stages of RA. We grouped 4047645 metagenomic genes into 6,224 KOs and 404 KEGG modules. We identified 12 KEGG modules that were significantly correlated with RA or OA ($q < 0.1$ or $q_{\text{partial}} < 0.1$, see online supplemental figure S2) and presented their variation across stages (figure 1A). We then used leave-one-out analysis to identify the MGS that most drove the correlations of these KEGG modules with RA or OA (figure 1B, online supplemental figure S3).

We found an evident decrease in glycosaminoglycan (CAG) metabolism across four RA stages and OA. It was mainly reflected by the significant decrease in K01197 (hyaluronoglucosaminidase) of dermatan sulfate (DS) degradation and the significant decrease in K10532 (heparan-alpha-glucosaminidase)

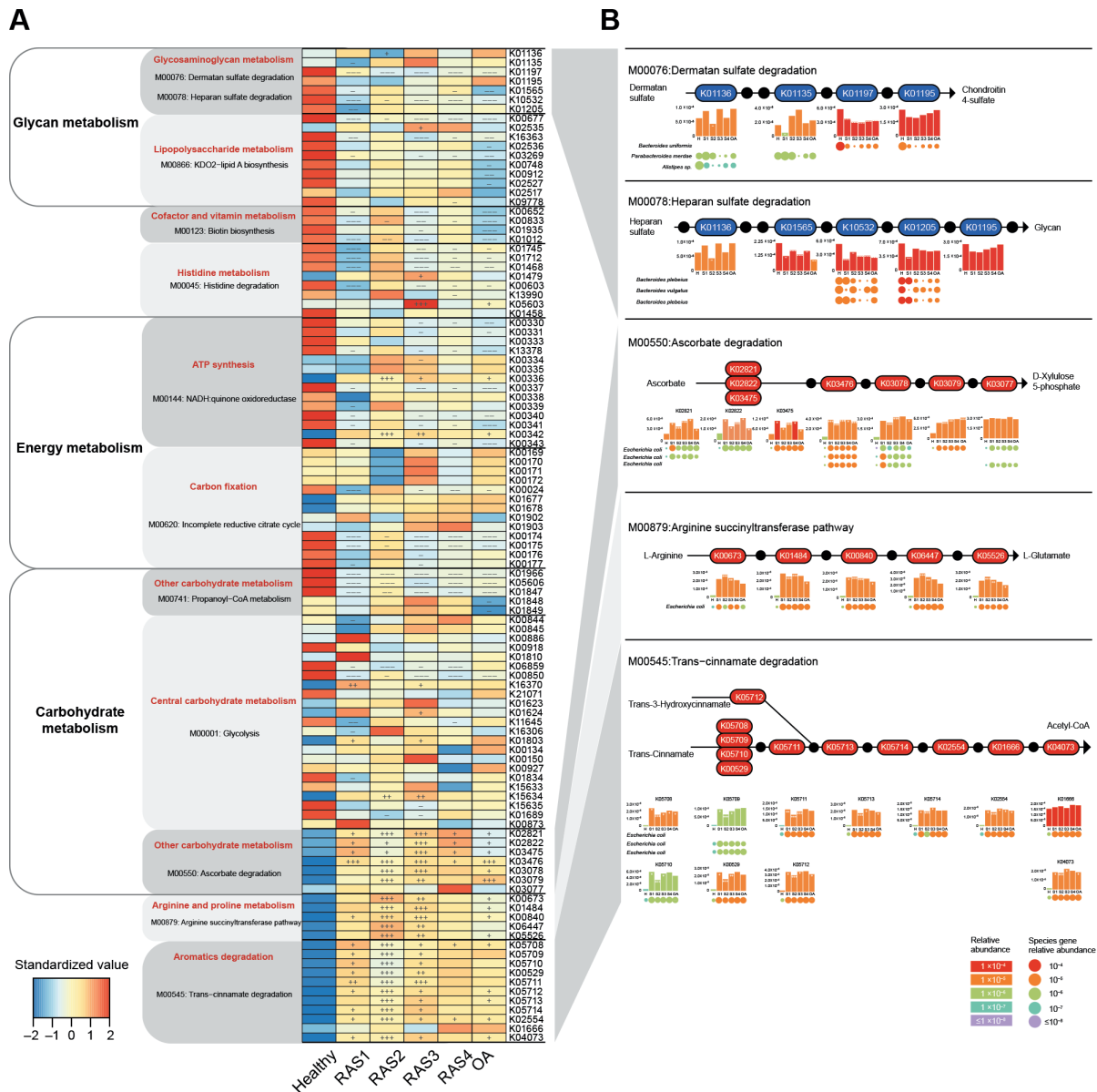


Figure 1 Stage-specific microbial functional profiles. Gene abundances were assessed for elevation or depletion in each of the arthritis stages, RAS1 (n=15), RAS2 (n=21), RAS3 (n=18), RAS4 (n=22) and OA (n=19) compared with the healthy individuals (n=27). (A) Relative abundance of KO genes in the KEGG modules that were significantly correlated with arthritis ($q < 0.1$ or $q_{\text{partial}} < 0.1$, see online supplemental figure S1). KO genes with a prevalence of 5% or higher are shown. (B) KO genes involved in specific KEGG pathway modules in (A) are shown in the KEGG pathway maps. Each box in a pathway represents a KO gene and is marked in red for elevation or in blue for depletion at any of the stages compared with healthy individuals. Bar plots show relative gene abundances averaged over samples within each of the five groups (healthy (H), RAS1 (S1), RAS2 (S2), RAS3 (S3), RAS4 (S4) and OA) and are coloured according to the values. Each KO gene is composed of MGS genes represented by circles. The sizes and colours of the circles are proportional to the relative abundances of the MGS genes. MGS genes are grouped into one row and indicated by the taxonomic name. The three MGS that most drove the correlation of the KEGG modules with arthritis types are shown. In all panels, significant changes are denoted as follows: +++, elevation with $p < 0.005$; ++, elevation with $p < 0.01$; +, elevation with $p < 0.05$; ---, depletion with $p < 0.005$; --, depletion with $p < 0.01$; -, depletion at $p < 0.05$; Mann-Whitney-Wilcoxon test. KEGG, Kyoto Encyclopaedia of Genes and Genomes; KO, KEGG ortholog; MGS, metagenomic species; OA, osteoarthritis.

N-acetyltransferase) of heparan sulfate (HS) degradation ($p < 0.05$, figure 1A). Chondroitin 4-sulfate is a major component of the extracellular matrix of many connective tissues, such as cartilage, bone and skin.¹² We found that the significant depletion of DS degradation would inhibit the production of chondroitin 4-sulfate (figure 1B), which might hurt the mechanical properties of the articular cartilage.¹² Moreover, the significant depletion of HS degradation might be a potential cause of the higher plasma level of HS observed in RA and OA

patients,^{20 21} which could promote arthritis progression by regulating protease activity.²² The most driving species of DS degradation and HS degradation were MGS *Bacteroides uniformis* and MGS *Bacteroides plebeius*, respectively. The genes of MGS *B. uniformis* related to K01197 were found most depleted in RAS2, while the genes of MGS *B. plebeius* related to K10532 were found most depleted in RAS3 and RAS4 (figure 1B). These results indicated that the depleted microbial function in DS degradation and HS degradation driven by *B. uniformis* and *B.*

plebeius, respectively, could promote RA and OA in a way of hurting articular cartilage.

We also identified elevated microbial functions that were related to inflammation such as the previously reported ascorbate degradation.⁷ Here, we found most of the KOs related to ascorbate degradation retained a higher level across RA stages and OA, especially in RAS2 and RAS3 ($p < 0.05$, figure 1A). Genes of K02821 (phosphotransferase system) in RAS1, K03475 (phosphotransferase system), K03476 (L-ascorbate 6-phosphate lactonase), and K03479 (L-ribose-5-phosphate 3-epimerase) were mostly driven by MGS *Escherichia coli* (figure 1B). The enhanced ascorbate degradation might contribute to the deficiency of the ascorbate reported in patients with RA²³ and were found positively correlated with multiple plasma cytokines ($q < 0.1$ or $q_{\text{partial}} < 0.1$, see online supplemental table S6), such as IL-1 β ($p = 5.44 \times 10^{-4}$), TNF- α ($p = 6.59 \times 10^{-4}$) and IL-6 ($p = 1.12 \times 10^{-3}$). Moreover, to confirm the effects of ascorbate on RA progression, we examined the plasma TNF- α level and IL-6 level, bone CT scans, and bone density of (1) normal DBA/1 mice, (2) DBA/1 mice with CIA and (3) DBA/1 mice with CIA and gavage of ascorbate. We found that the 3-month gavage of ascorbate to CIA mice can prevent the increase of TNF- α and IL-6 levels by half, inhibit bone destruction, and maintain bone density ($1.58 \pm 0.0034 \text{ g/cm}^3$), as compared with the CIA mice without ascorbate ($1.53 \pm 0.013 \text{ g/cm}^3$), and the normal group ($1.61 \pm 0.021 \text{ g/cm}^3$, see online supplemental figure S4).

For other elevated microbial functions, the trans-cinnamate degradation driven by MGS *E. coli*, where most KOs were notably elevated in RAS2, was also correlated with multiple cytokines ($q < 0.1$ or $q_{\text{partial}} < 0.1$, see online supplemental table S6), such as IL-13 ($p = 1.63 \times 10^{-5}$), IL-1 β ($p = 2.87 \times 10^{-5}$) and IL10 ($p = 4.10 \times 10^{-3}$). Moreover, the arginine succinyltransferase pathway driven by MGS *E. coli* was found significantly elevated mainly in RAS2 and RAS3 (figure 1). L-arginine is able to prevent bone loss induced by zinc oxide nanoparticles or by cyclosporin A, through anti-inflammatory mechanism²⁴ or nitric oxide production, respectively.²⁵ Both arginine succinyltransferase pathway and trans-cinnamate degradation was positively correlated with the elevation of rheumatoid factor ($p = 1.35 \times 10^{-3}$). Taken together, these results suggested that microbial dysfunction could promote RA progression mainly by hurting bone tissue and strengthening inflammation. The inflammation-related microbial dysfunction was extremely active in RAS2 and RAS3 and largely driven by *E. coli*.

Microbial invasion of the joint synovial fluid

Next, we investigated whether or in which stage microbial invasion of the joint synovial fluid happened. Enhanced gut permeability may render it possible for microbes and their products to translocate, triggering an immune response.^{6, 26} We thus speculated that gut microbes might invade the joint synovial fluid of patients with RA through the gut-joint axis. To test this, we performed 16S rRNA gene sequencing on the synovial fluid samples from another cohort of 271 patients in four RA stages, including RAS1 ($n = 52$), RAS2 ($n = 66$), RAS3 ($n = 67$) and RAS4 ($n = 86$). Notably, we were not able to obtain enough bacterial DNA for sequencing in samples of RAS1, RAS2 or RAS3, however, we could identify many microbes in samples of RAS4 (see online supplemental figure S5). We found that most of the microbes in joint synovial fluid were from phyla Proteobacteria and Firmicutes, and a total of 98 genera could also be detected in faecal metagenomic data (see online supplemental table S7). Moreover, we could recognise *E. lenta* and *B. longum* in most

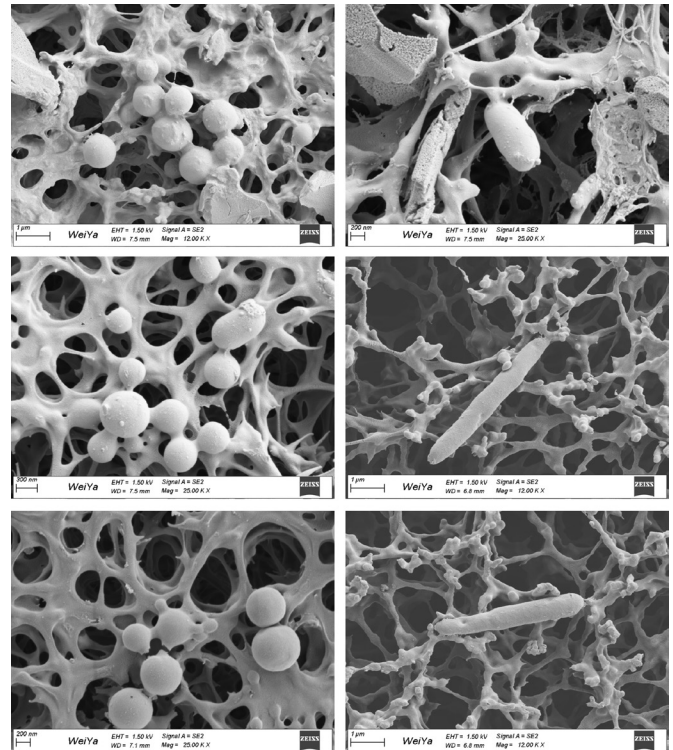


Figure 2 Scanning electron microscopy of the joint synovial fluid.

of the synovial fluid samples, both of which were observed to be exclusively elevated in faecal metagenome of patients in RAS4 from the multiomics cohort (see online supplemental table S4). In addition, *Prevotella copri* that has been reported highly correlated with RA^{8, 27} was also found abundant in most synovial fluid samples of patients in RAS4. We then randomly selected six synovial fluid samples per RA stage for bacteria isolation. Only from three synovial fluid samples of RAS4 can we separate bacteria. We then picked and sequenced three single colonies per synovial fluid sample. Five of the nine colonies were identified as *Clostridium sporogenes* strain, and three were identified as *Enterococcus gallinarum* strain, and one was identified as *Citrobacter freundii* strain (see online supplemental table S8). Interestingly, *Enterococcus gallinarum* and *Citrobacter freundii* could also be detected in faecal metagenomic data of 18% of patients with RA. We subsequently observed the corresponding synovial fluid samples using scanning electron microscopy, and found substances shaped like bacteria in rod-like or spherical forms (figure 2). Taken together, this multifaceted investigation has provided unprecedented evidence to support the existence of microbial invasion of the joints in the fourth stage of RA.

Stage-specific metabolomic profiles

We then introduced metabolomic data, and performed a random forest algorithm on 232 microbiome species, 6224 KOs and 277 metabolites to test their diagnostic potential for each stage of RA and OA (figure 3A–E). Metabolites exhibited the best area under the receiver operating characteristic curve (AUROC) in discriminating samples of four RA stages or OA from healthy samples, with AUROC ranging from 0.974 to 0.998. Other characteristics at the species and KO levels exhibited weaker discriminant ability, with AUROC ranging from 0.760 to 0.838 and from 0.799 to 0.852, respectively. The most prominent changes in metabolites were the significant increase of DL-lactate and gly-glu in RAS1 ($p = 2.15 \times 10^{-6}$, $p = 4.70 \times 10^{-4}$), the decrease of

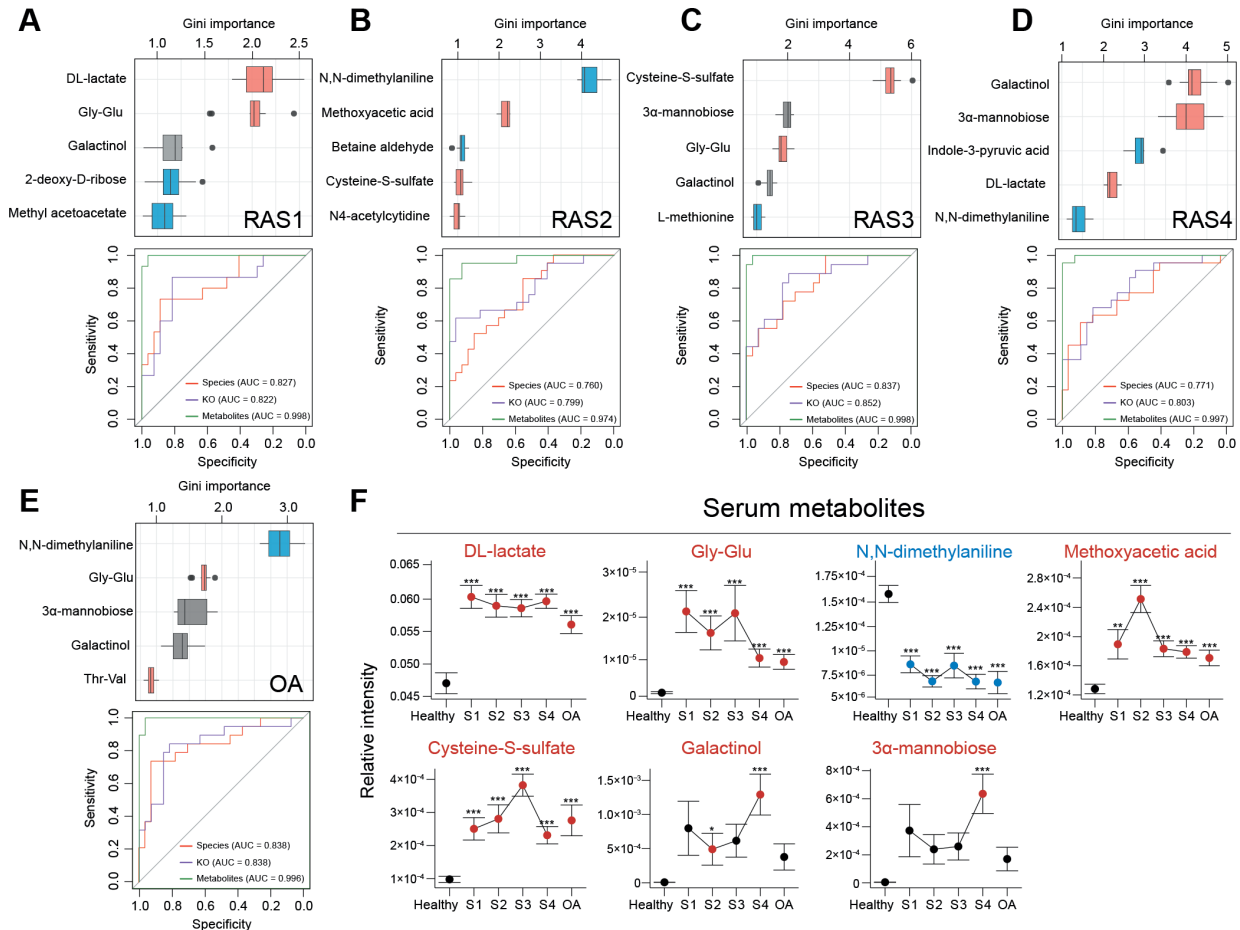


Figure 3 Multiomics diagnostic potential for the RA stage. A random forest algorithm was performed on 6224 KOs, 232 microbial species and 277 plasma metabolites in RAS1 (A), RAS2 (B), RAS3 (C), RAS4 (D) and OA (E). The Gini importance of the top five most discriminant metabolites are displayed. Boxes represent the IQR between the first and third quartiles and the line inside represents the median. Whiskers denote the lowest and highest values within the 1.5×IQR from the first and third quartiles, respectively. Boxes are marked in a specific colour to show the significant elevation ($p < 0.05$, red, Mann-Whitney-Wilcoxon test) or depletion ($p < 0.05$, blue, Mann-Whitney-Wilcoxon test) of the features in each of the arthritis stages compared with the healthy group. The ROC curves of the random forest model using microbial species, KOs, or metabolites were plotted, with AUC calculated by 10 randomised 10-fold cross-validation. The colour of the curve represents the category of the used features. (F) The dot plots show stage-specific abundance or concentration (mean±SE) of plasma metabolites, which are specified in (A–E). Four RA stages are connected to display the variance. Dots are coloured differently if the features are significantly elevated (red) or significantly depleted (blue), as compared with those of the healthy group. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$; Mann-Whitney-Wilcoxon test. AUC, area under curve; KEGG, Kyoto Encyclopaedia of Genes and Genomes; KO, KEGG ortholog; OA, osteoarthritis; RA, rheumatoid arthritis; ROC, receiver operating characteristic.

N,N-dimethylaniline and the increase of methoxyacetic acid in RAS2 ($p = 4.60 \times 10^{-8}$, $p = 1.28 \times 10^{-8}$), the increase of cysteine-S-sulfate ($p = 4.66 \times 10^{-12}$) in RAS3, the increase of galactinol and 3α-mannobiose in RAS4 ($p = 5.71 \times 10^{-5}$, $p = 5.68 \times 10^{-4}$), and the decrease of N,N-dimethylaniline and increase of gly-glu ($p = 2.00 \times 10^{-7}$, $p = 1.74 \times 10^{-3}$) in OA, as compared with a healthy state. The predominant metabolic disorders implicated a critical involvement in pathogenesis and a great diagnostic potential for RA stages.

Moreover, metabolic disorders could distinguish a given RA stage from not just healthy controls but also other RA stages or OA (figure 3F): Methoxyacetic acid in RAS2 ($p = 1.68 \times 10^{-4}$) or cysteine-S-sulfate in RAS3 ($p = 2.42 \times 10^{-4}$) or Galactinol and 3α-mannobiose in RAS4 ($p = 9.37 \times 10^{-3}$, $p = 4.89 \times 10^{-3}$), respectively, was higher than that in all the other RA stages and OA. Methoxyacetic acid was reported to have inhibitory effects on osteoblasts and could cause reductions in bone marrow cellularity.^{28–30} Additionally, cysteine-S-sulfate was a structural analogue of glutamate, acting as an agonist of

N-methyl-D-aspartate receptor (NMDA-R) whose expression and function in osteoclasts engaged in bone resorption.³¹ Therefore, notable elevations of methoxyacetic acid in RAS2 might hinder osteoblasts, whereas notable elevations of cysteine-S-sulfate in RAS3 might encourage osteoclasts. The imbalance between osteoblasts and osteoclasts would promote the bone erosion that occurred clinically in the third stage of RA. Moreover, DL-lactate in OA was less than that in all RA stages ($p = 0.037$), which might improve clinical differentiation of early RA from OA.

DISCUSSION

Our findings reveal dynamic shifts in gut microbiome and plasma metabolome, and their continuous roles in pathogenesis of RA across four successive stages (figure 4). Moreover, we demonstrate that microbial invasion of the joint synovial fluid happens in the fourth stage of RA.

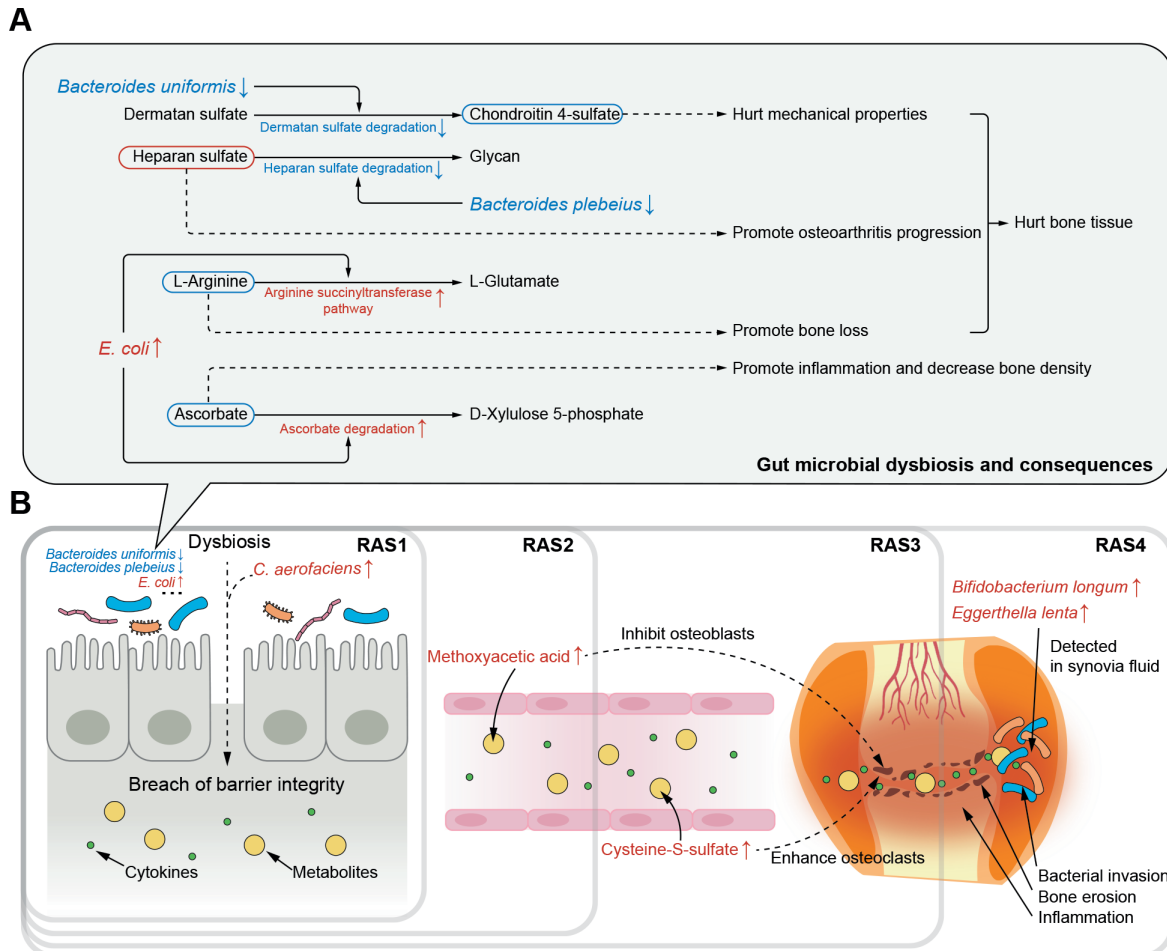


Figure 4 Potential pathogenesis across successive RA stages from multiomics perspective. (A) Potential mechanisms by which gut microbial dysbiosis play roles in RA pathogenesis through hurting bone tissue and increasing inflammation. The driving species, microbial dysfunction and related metabolites were extracted from figure 1B. The red or blue box of metabolites represents their speculated elevation or depletion according to the KEGG map. The dotted line represents the speculated effects of microbial and metabolic variation on arthritis pathogenesis. (B) The most representative effects of microbial dysbiosis and metabolic disorders on RA progression across successive stages. KEGG, Kyoto Encyclopaedia of Genes and Genomes; RA, rheumatoid arthritis.

The samples used for this study can fairly represent gut microbiome of each RA stage. Our hospital had tens of thousands of outpatients with arthritis per year and we have kept collecting samples from patients diagnosed with each stage of RA from 2017 to 2020. Considering the potential effects of clinical intervention on gut microbiome,³² in this study, we only recruited samples of those patients who had not received treatment within 1 month and were in the active period. Therefore, the microbial dysbiosis and metabolic disorders depicted here could serve as a profound reference for future studies in each stage of RA.

Clinical microbial intervention should take into account the stage of RA. We found each RA stage had its special elevated or depleted microbes that played a role in RA pathogenesis (figure 4A). Hence, it may not be adequate for clinical guidance to generally report microbial alterations in RA without information of the stage, as many studies have done.^{6–11} For instance, early inhibition of *C. aerofaciens* that was elevated exclusively in the first stage could help prevent increasing of gut permeability.⁹ Additionally, inhibition of *E. coli* in the second and third stage could help maintain the content of L-arginine that acted as an inhibitor of bone loss,^{24–25} as well as the content of anti-inflammatory ascorbate.³³ Moreover, certain species may need intervention across stages owing to its depletion during the

whole RA progression. A cross-stages restoration of *B. uniformis* could help maintain the content of chondroitin 4-sulfate to keep mechanical properties of the articular cartilage.¹²

Moreover, metabolic alterations kept considerable throughout RA progression, in spite of which we found that certain of these metabolites need a higher priority of intervention in a specific stage. In the second stage of RA, the aberrant elevation of methoxyacetic acid might have inhibitory effects on osteoblasts and cause reductions in bone marrow cellularity^{28–30} (figure 4B). The inhibited osteoblasts then drew the foreshadowing for the bone erosion that happened in the next stage. In the third stage, the considerable elevation of cysteine-S-sulfate might enhance the osteoclasts by NMDA-R interaction.³¹ The imbalance between osteoblasts and osteoclasts would then promote bone erosion that happened in the third stage and persisted in the late RA stages. Thus, methoxyacetic acid may be a targeted metabolite for treatment to patients in the second stage of RA and serve as a precaution against the upcoming third stage.

Our findings suggested that bacterial invasion of joint synovial fluid happened in the fourth stage of RA (figure 4B). Joint synovial fluid was generally considered sterile, and indeed, we failed to either extract enough DNA or isolate bacteria from the synovial fluid in the first three stages. However, in the

fourth stage, we succeeded to obtain bacterial 16S reads, isolate bacteria and observe substances shaped like bacteria in rod-like or spherical forms under scanning electron microscopy. Moreover, in the multiomics cohort, we found two faecal microbes elevated exclusively in the fourth stage of RA, *E. lenta* and *B. longum*, and their existence in the joint synovial fluid was validated by the other cohort. It might due to the buildup of the continuous damages in gut barrier and microbes and microbial metabolites would then be transferred to the joints via blood.⁶ Hence, for patients in the fourth stage of RA, in addition to routine medical therapies, specific treatments to the microbes in the joint synovial fluid may ameliorate the joint micro-environment to decrease synovial inflammation and inhibit potential bacterial effects.

This study also has limitations and prospects. First, a long-term follow-up investigation on a single individual throughout his/her RA development may reinforce the conclusions of this study. Second, it remains unclear how bacterial genetic materials are transferred from intestine to joint. It might be realised by bacteria transmission through blood or by means of extracellular vesicles or both. Third, the proposed links between microbial dysbiosis/metabolic disorders and RA can serve as a guidance for future experiments on RA pathogenesis. Lastly, additional researches into the synovial fluid microbiome and metabolome have the potential to reveal more sophisticated mechanisms underlying RA pathogenesis.

In conclusion, this study demonstrates microbial and metabolic roles in RA pathogenesis across four successive stages. A stage-specific intervention of microbial dysbiosis and metabolic disorders is warranted for prognosis and prevention of RA.

Author affiliations

¹First Affiliated Hospital of Shandong First Medical University, Institute of Medical Genomics, Biomedical Sciences College & Shandong Medicinal Biotechnology Centre, NHC Key Laboratory of Biotechnology Drugs (Shandong Academy of Medical Sciences), Key Lab for Rare & Uncommon Diseases of Shandong Province, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan, Shandong, China

²Key Laboratory of Molecular Biophysics of the Ministry of Education, Hubei Key Laboratory of Bioinformatics and Molecular-imaging, Center of AI Biology, Department of Bioinformatics and Systems Biology, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, Hubei, China

³Microbiome-X, National Institute of Health Data Science of China & Institute for Medical Dataology, Department of Biostatistics, School of Public Health, CheeLo College of Medicine, Shandong University, Jinan, Shandong, China

Contributors MC, YaZ, LZ, KN and JH designed the study, reviewed, and verified the data. MC, YaZ, YC, CZ, YuZ, SL, GC and ML collected samples and conducted experiments. MC, YaZ and KN conducted data analysis and produced the figures and tables. MC, YaZ, KN and JH wrote the manuscript. All authors revised the manuscript. MC, YaZ, LZ, KN and JH supervised the study. MC and YaZ are joint first authors. LZ, KN, and JH are joint senior authors. All authors approved the final version of the article. JH accepts full responsibility for the work and the conduct of the study, had access to the data, and controlled the decision to publish.

Funding This work was partially funded by the National Natural Science Foundation of China (grant numbers 31871334, 82003766, 32071465, and 31671374), the Academic Promotion Programme of Shandong First Medical University (grant number 2019LJ001) and the Key Research and Development project of Shandong Province (grant number 2021ZDSYS27).

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval The study was approved by the Ethics Committee of the First Affiliated Hospital of Shandong First Medical University (NO.2017-02 and NO.2020-011). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Whole-genome shotgun sequencing data are available in the Genome Sequence Archive (GSA) section of the National Genomics Data Center (project accession number CRA004348). 16S rRNA gene sequencing data are available in the Genome Sequence Archive (GSA) section of the National Genomics Data Center (project accession number CRA005811). Plasma metabolomic data are available in the MetaboLights (project accession number MTBLS5297).

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Mingyue Cheng <http://orcid.org/0000-0003-1243-5039>

Kang Ning <http://orcid.org/0000-0003-3325-5387>

Jinxiang Han <http://orcid.org/0000-0002-2507-9611>

REFERENCES

- Almutairi K, Nossent J, Preen D, *et al*. The global prevalence of rheumatoid arthritis: a meta-analysis based on a systematic review. *Rheumatol Int* 2021;41:863–77.
- Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. *Lancet* 2016;388:2023–38.
- Aletaha D, Neogi T, Silman AJ, *et al*. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League against rheumatism collaborative initiative. *Arthritis & Rheumatism* 2010;62:2569–81.
- Steinbrocker O, Traeger CH, Batterman RC. Therapeutic criteria in rheumatoid arthritis. *J Am Med Assoc* 1999;281:659–62.
- Barhum L. What does rheumatoid arthritis progression look like? 2021. Available: <https://www.verywellhealth.com/rheumatoid-arthritis-stages-of-progression-4768891>
- Zaiss MM, Joyce Wu H-J, Mauro D, *et al*. The gut-joint axis in rheumatoid arthritis. *Nat Rev Rheumatol* 2021;17:224–37.
- Zhao Y, Cheng M, Zou L, *et al*. Hidden link in gut-joint axis: gut microbes promote rheumatoid arthritis at early stage by enhancing ascorbate degradation. *Gut* 2022;71:1041–3.
- Pianta A, Arvikar S, Strle K, *et al*. Evidence of the immune relevance of *Prevotella copri*, a gut microbe, in patients with rheumatoid arthritis. *Arthritis Rheumatol* 2017;69:964–75.
- Chen J, Wright K, Davis JM, *et al*. An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis. *Genome Med* 2016;8:43.
- Smith PM, Howitt MR, Panikov N, *et al*. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* 2013;341:569–73.
- Maeda Y, Takeda K. Role of gut microbiota in rheumatoid arthritis. *J Clin Med* 2017;6:60.
- Henrotin Y, Mathy M, Sanchez C, *et al*. Chondroitin sulfate in the treatment of osteoarthritis: from in vitro studies to clinical recommendations. *Ther Adv Musculoskelet Dis* 2010;2:335–48.
- Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 2011;27:2957–63.
- de la Cuesta-Zuluaga J, Kelley ST, Chen Y, *et al*. Age- and sex-dependent patterns of gut microbial diversity in human adults. *mSystems* 2019;4:e00261–19.
- Lugli GA, Tarracchini C, Alessandri G, *et al*. Decoding the Genomic Variability among Members of the *Bifidobacterium dentium* Species. *Microorganisms* 2020;8. doi:10.3390/microorganisms8111720. [Epub ahead of print: 03 11 2020].
- Silvestre-Rangil J, Bagán L, Silvestre FJ, *et al*. Oral manifestations of rheumatoid arthritis. A cross-sectional study of 73 patients. *Clin Oral Investig* 2016;20:2575–80.
- Marriott D, Stark D, Harkness J. *Veillonella parvula* discitis and secondary bacteremia: a rare infection complicating endoscopy and colonoscopy? *J Clin Microbiol* 2007;45:672–4.
- Balakrishnan B, Luckey D, Taneja V. Autoimmunity-Associated gut commensals modulate gut permeability and immunity in humanized mice. *Mil Med* 2019;184:529–36.
- Tito RY, Cypers H, Joossens M, *et al*. Brief report: Dialister as a microbial marker of disease activity in spondyloarthritis. *Arthritis Rheumatol* 2017;69:114–21.
- Jura-Póltorak A, Komosinska-Vassev K, Kotulska A, *et al*. Alterations of plasma glycosaminoglycan profile in patients with rheumatoid arthritis in relation to disease activity. *Clin Chim Acta* 2014;433:20–7.

- 21 Shamdani S, Chantepie S, Flageollet C, *et al.* Heparan sulfate functions are altered in the osteoarthritic cartilage. *Arthritis Res Ther* 2020;22:283.
- 22 Severmann A-C, Jochmann K, Feller K, *et al.* An altered heparan sulfate structure in the articular cartilage protects against osteoarthritis. *Osteoarthritis Cartilage* 2020;28:977–87.
- 23 Abrams E, Sandson J. Effect of ascorbic acid on rheumatoid synovial fluid. *Ann Rheum Dis* 1964;23:295–9.
- 24 Abdelkarem HM, Fadda LH, El-Sayed EM, *et al.* Potential role of L-arginine and vitamin E against bone loss induced by nano-zinc oxide in rats. *J Diet Suppl* 2018;15:300–10.
- 25 Fiore CE, Pennisi P, Cutuli VM, *et al.* L-Arginine prevents bone loss and bone collagen breakdown in cyclosporin A-treated rats. *Eur J Pharmacol* 2000;408:323–6.
- 26 Manfredo Vieira S, Hiltensperger M, Kumar V, *et al.* Translocation of a gut pathobiont drives autoimmunity in mice and humans. *Science* 2018;359:1156–61.
- 27 Scher JU, Szczesnak A, Longman RS, *et al.* Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *Elife* 2013;2:e01202.
- 28 Brown NA, Holt D, Webb M. The teratogenicity of methoxyacetic acid in the rat. *Toxicol Lett* 1984;22:93–100.
- 29 Miller RR, Carreon RE, Young JT, *et al.* Toxicity of methoxyacetic acid in rats. *Fundam Appl Toxicol* 1982;2:158–60.
- 30 Sparks NR. The embryotoxic effects of harm reduction tobacco products on osteoblasts developing from human embryonic stem cells. University of California, Riverside 2018; chapter 4:73–104.
- 31 Li P, Sundh D, Ji B, *et al.* Metabolic Alterations in Older Women With Low Bone Mineral Density Supplemented With *Lactobacillus reuteri*. *JBMR Plus* 2021;5:e10478.
- 32 Schwartz DJ, Langdon AE, Dantas G. Understanding the impact of antibiotic perturbation on the human microbiome. *Genome Med* 2020;12:82.
- 33 Carr AC, Maggini S. Vitamin C and immune function. *Nutrients* 2017;9:1211.