Microbial dysbiosis and metabolic disorders promote rheumatoid arthritis progression across successive stages

Mingyue Cheng[1,2†], Yan Zhao[1†], Yazhou Cui[1], Yuguo Zha[2], Shufeng Li[1,3], Guangxiang Cao[1], Mian Li[1], Jinxiang Han[1*], Kang Ning[2*]

[1]First Affiliated Hospital of Shandong First Medical University, Biomedical Sciences College & Shandong Medicinal Biotechnology Centre, Shandong First Medical University & Shandong Academy of Medical Sciences, Jinan, China

[2]Key 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, China

[3]Shandong Key Laboratory of Rheumatic Disease and Translational Medicine, Department of Orthopedic Surgery, The First Affiliated Hospital of Shandong First Medical University & Shandong Provincial Qianfoshan Hospital, Jinan, China

[†]These authors contributed equally to this work.

[*]Correspondence:
Jinxiang Han (jxhan@sdfmu.edu.cn)
Kang Ning (ningkang@hust.edu.cn)

Keywords: Gut-joint axis, Rheumatoid arthritis, Microbiome, Metabolites, Dysbiosis, Pathogenesis, Successive stages, Osteoarthritis

Abstract

Rheumatoid arthritis (RA) is a chronic inflammatory disease that typically progresses through four stages, from mild stiffness to severe disability. The gut-joint axis is critical in RA pathogenesis, while its dynamic patterns and regulatory mechanisms throughout RA progression remain unknown. Here, we performed fecal metagenomic and serum metabolomic studies on a cohort of 122 patients from four distinct RA stages. Microbial dysbiosis and metabolomic disorders present stage-specific patterns and played crucial roles in RA pathogenesis across successive stages. In particular, *Escherichi coli* drove increased ascorbate degradation particularly in the second and third stages, promoting inflammation and bone destruction. Besides, abnormally high levels of serum methoxyacetic acid and cysteine-S-sulfate inhibited osteoblasts in the second stage and enhanced osteoclasts in the third stage, respectively. Furthermore, combined with analysis of 16S rRNA gene sequencing and scanning electron microscopy on joint synovial fluid from another cohort of 271 RA patients, microbial invasion of the joint was very likely to happen in the fourth stage. Our findings elucidate the role of microbial dysbiosis and metabolic disorders in successive stages of RA, which open up new avenues for RA prognosis and therapy.

Introduction

Rheumatoid arthritis (RA) affects over tens of millions of people worldwide[1]. RA is recognized clinically as a chronic, inflammatory, autoimmune disease that has four stages in progression[2,3]. In the first stage, the synovium of the joint is inflamed. Most people have minor symptoms such as stiffness upon awakening[2]. In the second stage, the inflamed synovium has caused damage to the joint cartilage. People begin to feel swelling, and have a restricted range of motion[2]. In the third stage, RA has proceeded to a severe state when bone erosion begins. By this stage, the cartilage on the surface of the bones has deteriorated, resulting in the bones rubbing against one another. People feel increased swelling and mobility problems[2]. In the fourth stage, certain joints were severely deformed and lost function. People suffer considerably from

swelling, stiffness, mobility loss, and disability[2]. 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, designated as gut-joint axis, via a range of mechanisms such as metabolic perturbation and immune response regulation[4-9]. For instance, increased abundance of *Prevotella* and *Collinsella* in RA patients are correlated with the production of T_H17 cell cytokines[10,11]. Additionally, mice models have confirmed microbial roles in RA. Reduction of *Bacteroides* spp. in mice with collagen-induced arthritis (CIA) may contribute to a local pro-inflammatory environment by inhibiting CD4+ T cell differentiation into regulatory T cells[12]. Microbial metabolites have also been correlated to immunity regulation in RA. Short-chain fatty acids (SCFAs) have a local protective effect on colonic inflammation, via binding to FFAR2 on colonic T cells and ILC3s[8,13]. SCFA administration to mice with CIA can reduce the severity of arthritis by modulation of IL-10[5,8]. Comprehensive metagenomic and metabolomic analyses may therefore enhance our understanding about the gut-joint axis. The role of the gut-joint axis across successive stages of RA is understudied[5,14], where more examination may provide an alternative approach to ameliorate RA progression.

In this study, we performed fecal metagenomic and serum metabolomic studies on four stages of RA to depict the stage-specific patterns of microbial dysbiosis and metabolic disorders. Our study demonstrates that microbial dysbiosis and metabolic disorders plays continuous roles in RA pathogenesis across successive stages. In particular, the depletion of *Bacteroides uniformis* and *Bacteroides plebeius* drove microbial dysfunction in glycosaminoglycan metabolism, continuously hurting articular cartilage. The elevation of *Escherichi coli* drove microbial dysfunction mostly in the second and third stages, promoting inflammation and bone lose. Moreover, metabolome was extremely disordered in the whole progression and induced bone erosion manly through the abnormal elevation of methoxyacetic acid in the second stage and cysteine-S-sulfate in the third stage. Furthermore, together with analysis of 16S rRNA gene sequencing and scanning electron microscopy on joint

synovial fluid from another cohort of 271 RA patients, we concluded that the buildup of these effects might lead to microbial invasion of the joint in the fourth stage of RA: Gut microbes *Eggerthella lenta* and *Bifidobacterium longum* were elevated exclusively in the fourth stage, and they could also be detected in the joint synovial fluid of patients in the fourth stage. Collectively, these findings may open new opportunities for prognosis, pain relief, and therapies throughout RA progression.

Results

Multi-omics profiled cohort includes healthy controls and patients with rheumatoid arthritis or osteoarthritis

To characterize the gut-joint axis across successive stages of arthritis, we analyzed gut metagenomic data and serum metabolomic data from 27 healthy controls, 76 rheumatoid arthritis (RA) patients, 19 osteoarthritis (OA) patients. RA samples were grouped into four RA stages including RA stage 1 (RAS1, n = 15), RA stage 2 (RAS2, n = 21), RA stage 3 (RAS3, n = 18), and RA stage 4 (RAS4, n = 22) according to the rheumatoid diagnostic score[15] (**Supplementary Tables 1**). We employed two separate processes for gut metagenomic data analysis: 1) One was based on 4 million predicted genes and investigated the stage-specific functional composition in arthritis progression, as well as the metagenomic species (MGS) that most drove the correlation of these microbial functions with arthritis[16] (**Fig. 1a**); and 2) The other reported the 232 classified microbial species composition across arthritis stages[17], profiled by MetaPhlAn2[18] (**Fig. 1b**). We also explored the variance in serum metabolomic across RA stages (**Fig. 1c**). Furthermore, we identified the most notable variation of features at the multi-omics level that could identify each arthritic stage from the healthy condition with high accuracy using the random forest (RF) algorithm on microbial species, microbial KOs, and serum metabolites, respectively (**Fig. 1d**). Finally, we attempted to determine the existence of microbial invasion of the joint in the fourth stage of RA.

Patients with rheumatoid arthritis or osteoarthritis have altered microbial composition compared to healthy controls

We firstly identified arthritis-related microbial species that were substantially elevated or depleted in RA or OA groups, when compared with the healthy group ($P < 0.05$, $FDR < 0.05$, Mann-Whitney-Wilcoxon test, **Supplementary Tables 2**). The union set of these arthritis-related species ($n = 15$) was then used for linear discriminant analysis (LDA) to separate arthritis samples from healthy samples (**Fig. 2a**). We discovered that arthritis gut microbiota could be discriminated from healthy microbiota, while we did not find an evident separation between RA and OA samples in LDA. The most abundant two arthritis-related species were *Bacteroides uniformis* whose proportion was mainly contributed by healthy samples, and *Escherichia. coli* whose proportion was mainly contributed by RA and OA samples (**Fig. 2b**). Moreover, the RA group contributed dominantly to the proportion of *Clostridium hathewayi*, compared to its extremely low level in healthy and OA samples (**Fig. 2b,c)**. Thus, the significant elevation of *C. hathewayi* might be the RA-specific microbial variation that could discriminate RA samples from OA samples.

Microbial species present stage-specific alterations throughout the progression of rheumatoid arthritis

Next, we sought to explore the stage-specific taxonomic composition throughout the RA progression. We examined 232 classified species and compared their abundances in each of RA stages and OA to those in the healthy state (**Supplementary Tables 3**). The elevated species in RA progression were mostly from the phyla Firmicutes and Actinobacteria, while the depleted species were predominantly from the phylum Bacteroides (**Fig. 2d**). Additionally, we exhibited the stage-specific abundance of the identified 15 arthritis-related species, as well as species who or whose genera were previously recognized as arthritis-related in other studies[19] (**Fig. 2d**).

For the arthritis-depleted species (**Fig. 2d**), we found that *B. uniformis* (RAS1: $P = 1.4 \times 10^{-3}$, RAS2: $P = 5.0 \times 10^{-3}$, RAS3: $P = 7.2 \times 10^{-3}$, RAS4: $P = 5.7 \times 10^{-3}$) and *C. fastidiosus* (RAS1: $P = 0.041$, RAS2: $P = 0.024$, RAS3: $P = 0.022$) were significantly depleted only during the RA progression but not in OA, whereas *B. cellulosilyticus* was significantly depleted in OA ($P = 4.72 \times 10^{-4}$), and in the very early and late stages of the RA progression (RAS1: $P = 0.023$, RAS4: $P = 6.24 \times 10^{-3}$).

For the arthritis-elevated species (**Fig. 2d**), we found that *C. hathewayi* was not significantly elevated in the very early RA stage, though kept a significantly higher level during the subsequent RA stages than that in the healthy state (RAS2: $P = 2.52 \times 10^{-3}$, RAS3: $P = 5.92 \times 10^{-3}$, RAS4: $P = 2.96 \times 10^{-3}$). In addition, *Bifidobacterium dentium* also maintained considerable increased abundance from RAS2 to RAS4 (RAS2: $P = 7.16 \times 10^{-3}$, RAS3: $P = 3.70 \times 10^{-3}$, RAS4: $P = 9.15 \times 10^{-4}$). *B. dentium* was reported to be associated with the development of dental caries and periodontal disease, both of which were particularly prevalent in patients with RA[20,21]. *Streptococcus anginosus* and *S. parasanguinis* kept significant elevation ($P < 0.05$) throughout the RA progression (**Fig. 2d**), while a few of other species from genus *Streptococcus* were elevated stage-specifically. *S. mitis* ($P = 1.21 \times 10^{-3}$) and *S. sanguinis* ($P = 5.37 \times 10^{-4}$) were mainly elevated in RAS1, while *S. salivarius* was mainly elevated in RAS3 ($P = 0.035$) and RAS4 ($P = 0.022$). These results suggest that although a few microorganisms were identified as correlated with arthritis, they might not all keep abundant or depleted throughout the progression. Part of them only preferred certain stages.

Certain microbial species are only significantly altered in a single stage of rheumatoid arthritis

Besides those varied species shared by two or more arthritis stages, we also cared about those species that were only elevated or depleted at one specific arthritis stage when compared to the healthy state ($P < 0.05$, Mann-Whitney-Wilcoxon test, **Supplementary Tables 4**). *Collinsella aerofaciens*, for example, was found

significantly elevated only in RAS1 (P = 0.043). *C. aerofaciens* was previously reported to generate severe arthritis when inoculated into collagen-induced arthritis (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[11]. Therefore, the significant elevation of *C. aerofaciens* might contribute to the early breach in gut barrier integrity during RA progression[5]. The translocation of microbial products through the breach would then trigger the subsequent clinical arthritis[5].

Veillonella parvula, whose infection could cause osteomyelitis[22], 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. Additionally, mice gavaged with *E. lenta* were reported to show an increased gut permeability and produce proinflammatory cytokines IL-6, IL-21, IL-23[23]. *Dialister invisus* was found significantly elevated only in OA microbiota (P = 0.041). *D. invisus* was previously reported to be more abundant in the inflamed ileal and colonic biopsy from patients with spondyloarthritis than in the noninflamed biopsy, and was positively correlated with spondyloarthritis severity[24]. We also found that the species *Bacteroides ovatus*, *Bacteroides coprocola*, and *Bacteroides faecis* were depleted significantly only in RAS1 ($P = 4.2 \times 10^{-3}$), RAS2 (P = 0.036), and RAS4 (P = 0.020), respectively. These results suggest that certain microbial species in arthritis patients present more abundant or more depleted exclusively in a particular arthritis stage, for implementing specific functions such as increasing gut permeability and inducing proinflammatory cytokines.

B. uniformis- and _B. plebeius_-driven microbial dysfunction in glycosaminoglycan metabolism promotes arthritis progression by hurting articular cartilage

To understand the variation of microbial functions across arthritis stages, we grouped 4,047,645 metagenomic genes into 6224 KEGG orthologs (KOs) and 404 KEGG modules. We identified 12 KEGG modules that were significantly correlated with

arthritis ($P < 0.05$, FDR < 0.1, Mann-Whitney-Wilcoxon test, **Fig. 3a**, **Extended Data Fig. 1**, **Supplementary Tables 5** and **6**). We then compared KOs' abundances within these arthritis-related KEGG modules between arthritis stages and healthy state (**Fig. 3a** and **Supplementary Tables 7**), and used leave-one-out analysis to identify the metagenomic species (MGS) that most drove the correlations of these KEGG modules with arthritis (**Fig. 3b** and **Extended Data Fig. 2**).

We found an evident decrease in glycosaminoglycan (CAG) metabolism in all the arthritis stages. It was mainly reflected by the significant decrease in K01197 (hyaluronoglucosaminidase, RAS1: $P = 2.31 \times 10^{-3}$, RAS2: $P = 7.78 \times 10^{-5}$, RAS3: $P = 2.00 \times 10^{-5}$, RAS4: $P = 6.70 \times 10^{-4}$, OA: 8.65×10^{-4}) of dermatan sulfate (DS) degradation and the significant decrease in K10532 (heparan-alpha-glucosaminide N-acetyltransferase, RAS1: $P = 2.30 \times 10^{-4}$, RAS2: $P = 0.045$, RAS3: $P = 3.2 \times 10^{-3}$, RAS4: $P = 4.66 \times 10^{-3}$, OA: $P = 3.72 \times 10^{-4}$) of heparan sulfate (HS) degradation (**Fig. 3a**). The compressive features of the articular cartilage are determined by CAGs, predominantly chondroitin sulfate (CS) and hyaluronic acid[25]. CS is a major component of the extracellular matrix of many connective tissues, including cartilage, bone, skin, ligaments, and tendons[26]. We found that the significant depletion of DS degradation at all the arthritis stages would decrease the production of chondroitin 4-sulfate (**Fig. 3b**), which hurt the mechanical properties of articular cartilage of patients with arthritis. Additionally, total CAGs, CS, and heparan sulfate/heparin (HS/H) were reported to be higher in plasma of RA patients than in plasma of healthy subjects[27]. Moreover, the sulfation of HS was reported to be increased in cartilages of patients with OA[28], and a decreased HS content or a reduced sulfation level was reported to protect against OA progression by regulating protease activity[25]. Thus, the significant depletion of HS degradation in microbial function would prevent HS from being degraded, promoting arthritis progression (**Fig. 3b**). The most driving species of dermatan sulfate degradation and heparan sulfate 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 (**Fig. 3B**). 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.

***E. coli*-driven microbial dysfunction promotes inflammation in arthritis pathogenesis particularly in the second and third stages of RA**

We also identified certain arthritis-elevated microbial functions that were related to pro-inflammation. As we have previously reported, *Escherichia coli* could promote arthritis progression by enhancing ascorbate degradation[9]. Ascorbate could prevent the development of inflammatory arthritis by facilitating collagen synthesis, moderating autoimmune responses, and ameliorating inflammation[29,30]. Here, we found most of the KOs related to ascorbate degradation kept a higher level in RA and OA progression, especially in RAS2 and RAS3 (**Fig. 3a**). MGS *E. coli* drove the correlations of ascorbate degradation with arthritis. Genes of K02821 (phosphotransferase system) in RAS1, K03475 (phosphotransferase system), K03476 (L-ascorbate 6-phosphate lactonase), and K03479 (L-ribulose-5-phosphate 3-epimerase) in all arthritis stages were mostly contributed by MGS *E. coli* (**Fig. 3b**). Hence, the deficiency of ascorbate reported in patients with RA[31], might be related to the enhanced ascorbate degradation in gut microbial dysfunction. Moreover, we found that the enhanced ascorbate degradation was evidently correlated with the elevation of multi serum cytokines, including IL-1β ($P = 5.44 \times 10^{-4}$, q = 0.037), IL-17 ($P = 6.59 \times 10^{-4}$, q = 0.050), IL-8 ($P = 7.45 \times 10^{-4}$, q = 0.07), IL-2 ($P = 9.37 \times 10^{-4}$, q = 0.064), IL-6 ($P = 1.12 \times 10^{-3}$, q = 0.061), IL-12p70 ($P = 3.47 \times 10^{-3}$, q = 0.093), IL-4 ($P = 4.2 \times 10^{-3}$, q = 0.097), TNF-α ($P = 6.59 \times 10^{-3}$, q = 0.050).

To confirm the effects of ascorbate on RA progression, we examined the serum 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 three-month gavage of ascorbate can prevent the

increase of TNF-α and IL-6 levels, inhibit bone destruction and bone loss (**Extended Data Fig. 3**).

For other arthritis-elevated microbial functions, the arginine succinyltransferase pathway was found significantly elevated mainly in RAS2 and RAS3 (**Fig. 3a**). L-arginine has been reported to be able to prevent bone loss induced by zinc oxide nanoparticles or by cyclosporin A, through anti-inflammatory mechanism[32] or nitric oxide production, respectively[33]. The most driven MGS of the arginine succinyltransferase pathway was MGS *E. coli* (**Fig. 3b**). Moreover, another KEGG module named trans-cinnamate degradation was also most driven by MGS *E. coli*, whose most KOs were significantly elevated in RA and OA, especially in RAS2 (**Fig. 3a,b**). Trans-cinnamate degradation was positively correlated with the elevation of multi serum cytokines, including IL13 ($P = 1.63 \times 10^{-5}$, $q = 4.41 \times 10^{-3}$), IL-1β ($P = 2.87 \times 10^{-5}$, $q = 3.90 \times 10^{-3}$), IL10 ($P = 4.10 \times 10^{-5}$, $q = 5.57 \times 10^{-3}$), TNF-α ($P = 6.97 \times 10^{-5}$, $q = 9.47 \times 10^{-3}$), IL2 ($P = 7.40 \times 10^{-5}$, $q = 0.010$), IL-8 ($P = 2.46 \times 10^{-4}$, $q = 0.067$), IFNγ ($P = 3.92 \times 10^{-4}$, $q = 0.036$), IL4 ($P = 1.78 \times 10^{-3}$, $q = 0.061$), and IL-12p70 ($P = 2.61 \times 10^{-3}$, $q = 0.089$). Taken together, these results suggested that arthritis-related microbial functional change could promote RA and OA progression by hurting bone tissue and strengthening inflammation. The inflammation-related microbial function was extremely active in RAS2 and RAS3 and largely driven by *E. coli*.

Gut microorganisms may invade joint synovial fluid in the fourth stage of RA

To examine whether and in which of stages microorganisms had invaded joint synovial fluid (SF) of patients with RA, we collected samples of joint SF from another cohort of 271 patients in four distinct stages of RA, including RAS1 (n = 52), RAS2 (n= 66), RAS3 (n = 67), and RAS4 (n = 86). Enhanced gut permeability may render it possible for microorganisms and their products to translocate, triggering an immune response[5,34]. Additionally, a few studies have shown the presence of microbial DNA in the joint SF[35-38]. We thus speculated that gut microorganisms might

invade the joint SF of patients with RA through the gut-joint axis. To test this speculation, we next perform 16S rRNA gene sequencing on the SF samples. Notably, we were not able to obtain enough bacterial DNA for sequencing in samples of RAS1, RAS2, or RAS3, however, we could identify many microorganisms in samples of joint SF from patients in RAS4 (**Extended Data Fig. 4**). We found that most of the joint-SF microorganisms were from phyla Proteobacteria and Firmicutes (**Extended Data Fig. 4a**). Moreover, we could recognize *Eggerthella lenta* and *Bifidobacterium longum* in most of the SF samples (**Extended Data Fig. 4b**), both of which were observed to be exclusively elevated in the intestine of patients in RAS4 (**Supplementary Tables 4**). In addition, *Prevotella copri* that has been reported highly correlated with RA[10,39] was also found abundant in most SF samples of patients in RAS4.

We next randomly selected six SF samples per RA stage for bacteria isolation. Only from three SF samples of RAS4 can we separate bacteria (**Extended Data Fig. 5**). We then picked and sequenced three single colonies per SF sample. Five of the nine colonies were identified as *Clostridium sp.*, three were identified as *Enterococcus sp.*, and one was identified as *Citrobacter sp.* (**Supplementary Tables 8**). We subsequently observed the corresponding SF samples using scanning electron microscopy, and found substances shaped like bacteria in rod-like or spherical forms (**Extended Data Fig. 6**). Taken together, this multi-faceted investigation has provided unprecedented evidence to support the existence of microbial invasion of the joint in the fourth stage of RA.

Rheumatoid arthritis has stage-specific biomarkers indicating both microbial dysbiosis and metabolic disorders

Microbial dysbiosis has exhibited stage-specific patterns, playing a crucial role in successive stages of RA, via microbial dysfunction. It was also speculated that dysbiosis might promote RA progression via metabolic perturbation[5].

To determine the most discriminant characteristics of each arthritis stage at the multi-omics level, we performed an RF algorithm on 232 microbiome species, 6224 KOs, and 277 metabolites (**Fig. 4**), and displayed the most notable microbial and metabolic features of each stage in **Fig. 5a**. Metabolites exhibited the best AUROC in discriminating samples of all the arthritis stages from healthy samples, with AUROC ranging from 0.974 to 0.998. Other characteristics at the species and KO levels exhibited weaker discriminant ability than features at the metabolite level, with AUROC ranging from 0.760 to 0.838 and 0.799 to 0.852, respectively (**Fig. 4**). Thus, metabolic disorders were predominant in each stage of arthritis, implicating a critical involvement in pathogenesis throughout the disease development.

Abnormal elevation of serum lactate occurrs in the early stage of RA and is correlated with inflammation

Each of the arthritis stages had the respective most prominent change in metabolites (**Fig. 4** and **Fig. 5a**), even though all of them were strongly different from those of health status (**Fig. 5b**). 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 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 to a healthy state. A considerable increase in lactate, which was the most noticeable alteration in metabolites in RAS1 (**Fig. 4a**), occurred early in RA development and remained throughout the subsequent RA stages and OA (**Fig. 5b**). Lactate in the inflamed tissue was reported to promote the upregulation of the lactate transporter SLC5A12 by human CD4+ T cells[40]. This upregulation led to increased IL17 production via nuclear PKM2/STAT3 and enhanced fatty acid synthesis[40]. In addition, antibody-mediated blockade of SLC5A12 could ameliorate the severity of arthritis in a mouse model[40]. We found serum DL-lactate was positively correlated with a certain number of

inflammatory cytokines such as IL-6 (r = 0.33, $P = 1.86 \times 10^{-5}$, $q = 3.69 \times 10^{-4}$), IL-17 (r = 0.32, $P = 4.23 \times 10^{-5}$, $q = 9.76 \times 10^{-4}$) and TNF-alpha (r = 0.40, $P = 1.84 \times 10^{-7}$, $q = 1.28 \times 10^{-5}$). As the end product of glycolysis[41], the lactate production would be affected by the intervention of glycolysis, such as the disruption of the glucose-6-phosphate isomerase (GPI, K06859) and phosphofructokinase (PFK, K00850). Both GPI and PFK, as the upstream glycolytic enzyme, were found significantly decreased in RA stages ($P < 0.05$, $q < 0.1$), while GPI was not significantly decreased in the OA stage (**Fig. 3a**). As a consequence, microbial functional variation in glycolysis might result in an increased concentration of DL-lactate at RA stages compared to OA stages and healthy state, contributing to the elevated inflammatory cytokines in RA stages.

Metabolic disorders break the balance between osteoblasts and osteoclasts in the medium stage of rheumatoid arthritis

Furthermore, we discovered that a few noticeable alterations in metabolites might distinguish a given arthritis stage from not just a healthy state but also other arthritis stages. (**Fig. 5b**): (1) DL-lactate in OA was less than that at all RA stages ($P = 0.037$); (2) 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 at all the other RA stages and OA. Such dramatic changes in arthritis serum metabolites suggested that they might be more than just a metabolomic depiction of each stage of arthritis; they could also be potential promoters for the next stage. For instance, methoxyacetic acid was reported to have inhibitory effects on osteoblasts and could cause reductions in bone marrow cellularity[42-44]. The dominant elevation of methoxyacetic acid in RAS2 might be a precursor to bone erosion in RAS3. Moreover, the notable elevation of cysteine-S-sulfate in RAS3 might promote the osteoclasts that occurred in the late RA stages. cysteine-S-sulfate, a structural analog of glutamate, acted as an agonist of N-methyl-D-aspartate receptor (NMDA-R) whose expression and function in osteoclasts engaged in bone resorption[45]. Taken together, aberrant elevations of

methoxyacetic acid in RAS2 hindered osteoblasts, whereas abnormal elevations of cysteine-S-sulfate in RAS3 encouraged osteoclasts. As a result of the imbalance between osteoblasts and osteoclasts, bone degradation occurred in the late stages of RA.

Discussion

We performed metagenomic and metabolomic studies on a cohort of 122 RA patients, together with 16S rRNA gene sequencing and scanning electron microscopy on joint synovial fluid samples from another cohort of 271 RA patients, and drew several important conclusions. First, we found that microbial dysbiosis and metabolic disorders presented dynamic patterns across successive stages of RA. Second, these dysbiosis and disorders continuously promoted RA progression by increasing gut permeability, inducing inflammation, and hurting bone tissue. Microbial dysfunction in dermatan sulfate and heparan sulfate degradation continuously hurting the articular cartilage across successive stages of RA, as well as in OA. Third, certain of these influences were more concentrated on a particular stage of RA. Microbial dysfunction related to inflammation was focused on the second and third stages and was mostly driven by *E. coli*. Metabolic disorders induced bone erosion through the abnormal elevation of methoxyacetic acid in the second stage and cysteine-S-sulfate in the third stage. Finally, there was strong evidence that microbial invasion of the joint happened in the fourth stage of RA. Synovial-fluid 16S rRNA gene sequencing was capable of detecting *Eggerthella lenta* and *Bifidobacterium longum*, both of which were elevated in fecal metagenomics of patients, exclusively in the fourth stage of RA rather than other stages. Moreover, *Clostridium sp.*, *Enterococcus sp.*, and *Citrobacter sp.* were separated from the joint synovial fluid. Under scanning electron microscopy, we could observe substances shaped like bacteria in rod-like or spherical forms. However, we were unable to detect bacteria in the joint synovial fluid from patients in other stages of RA due to the low DNA content. We conclude that the buildup of microbial dysbiosis and metabolic disorders across successive RA stages may induce the microbial invasion of the joint in the last stage of RA.

The findings of the continuous effects of microbial dysbiosis and metabolic disorders across successive RA stages could be of considerable clinical importance (**Fig. 6**). First, hurting bone tissue (**Fig. 6a**). Depleted *B. uniformis* drove the decreased dermatan sulfate degradation, leading to less production of chondroitin 4-sulfate. The low content of chondroitin 4-sulfate would inhibit the formation of articular cartilage[25,26]. Additionally, depleted *B. plebeius* drove the less degradation of HS. A high HS content would promote OA progression[25,28]. Second, increasing inflammation particularly in the second and the third of RA (**Fig. 6b**). Elevated *E. coli* drove enhanced degradation of L-Arginine through the arginine succinyltransferanse pathway. L-Arginine acted as an inhibitor of bone loss via anti-inflammation[32] and nitric oxide production[33]. Besides, elevated *E. coli* drove enhanced ascorbate degradation. Ascorbate is an anti-inflammatory vitamin[29,30]. We found that three months of ascorbate treatment to mice with CIA could relieve arthritis, by preventing bone destruction and inflammation. Owing to the sequencing limitation in this study, we are unable to pinpoint the strain of microbial species in these gut-joint axes. Further explorations are required for determining strain identity, maybe through third-generation sequencing, as well as for testing its effects in a sterile mice model. In clinical practice, the intervention based on these proposed mechanisms may profoundly improve the therapy for patients during the whole RA progression, through restoring *B. uniformis* and *B. plebeius*, and inhibiting *E. coli*, or supplying patients with ascorbate and chondroitin 4-sulfate.

Apart from clinical treatments for microbial dysbiosis and metabolic disorders in the whole RA progression, specific treatments are necessary for each RA stage to not only relieve the symptoms in the current stage but also inhibit the onset of the subsequent stages. Our findings suggested that certain effects of microbial dysbiosis or metabolic disorders on RA progression were more focused on a particular stage. In the early stage of RA, elevated *C. aerofaciens* and depleted *Parabacteroides*. *distasonis* both contributed to the increased gut permeability (**Fig. 6b**). *C. aerofaciens* was found elevated exclusively in the first stage of RA, rather than other stages. *C.*

aerofaciens was reported to induce expression of IL-17 network cytokines and increase gut permeability via reducing the expression of tight junction protein in epithelial cells[11]. *P. distasonis* was depleted in most of RA stages and OA with a remarkable ability to distinguish arthritis samples from healthy samples, particularly in the first stage (**Fig. 5b**). Previous studies have reported the anti-inflammation effect of *P. distasonis* such as its correlation with CRP[46], TNF-α[47], and IL-1β[47], and that the oral administration of its antigens could reduce intestinal inflammation in colitis mice[48]. However, the role of *P. distasonis* in RA pathogenesis remained unknown. A recent study found that *P. distasonis* could preserve the intestinal barrier integrity in azoxymethane-treated A/J mice, and possess anti-inflammatory properties and anti-tumor properties through the inhibition of signaling via TLR4, MYD88, and Akt, as well as the stimulation of apoptosis[49]. We thus speculate that the considerable depletion of *P. distasonis* occurs in the very early of RA, contributing to inflammation and the collapse of intestinal barrier integrity, though this possible mechanism warrants further investigations. Intervention on *C. aerofaciens* and *P. distasonis* therefore may have a higher priority in the first stage of RA than in subsequent stages to inhibit the early increase of gut permeability and inflammation.

In the second stage of RA, the aberrant elevation of methoxyacetic acid had inhibitory effects on osteoblasts and caused reductions in bone marrow cellularity[42-44] (**Fig. 6b**). The inhibited osteoblasts drew the foreshadowing for the bone erosion that happened in the next stage. In the third stage, the considerable elevation of cysteine-S-sulfate enhanced the osteoclasts by NMDA-R interaction[45] (**Fig. 6b**). The imbalance between osteoblasts and osteoclasts then resulted in bone erosion that happened in the third stage and persisted in the late RA stages (**Fig. 6b**). 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. In the fourth stage of RA, continuous damages in gut permeability, intestinal immunity, and bone tissue resulted in the extremely high elevation of proinflammatory cytokines and severe bone erosion, and even microbial invasion of the joint. Due to the collapse of the intestinal barrier,

microbes and microbial metabolites might be transmitted to the joint via blood[5]. In addition to routine medical therapies for immunity regulation, specific treatments to the microbes in the joint synovial fluid of the patients in the fourth stage of RA may ameliorate the joint micro-environment to decrease synovial inflammation and prevent potential bacterial effects.

This study also has prospects and limitations. Firstly, it may aid in enhancing RA therapy across populations by assessing whether RA patients with different genetic backgrounds exhibit distinct gut-joint links across stages. Secondly, a long-term follow-up investigation on a single individual throughout his/her RA development may reinforce the conclusions of this study. Lastly, additional researches into the joint-SF microbiome and metabolome have the potential to figure out more sophisticated mechanisms underlying RA pathogenesis.

In conclusion, we demonstrate dynamic shifts in gut microbiome and serum metabolome, and their continuous roles in pathogenesis of RA across four successive stages. Microbial dysbiosis and metabolic disorders promotes RA through increasing gut permeability, triggering inflammation, and hurting bone tissue. Although these effects persist in the whole RA progression, certain of them are more focused on specific stages. Furthermore, the buildup of the effects of these dysbiosis and disorders may lead to the collapse in the fourth stage of RA, as represented by the microbial invasion of the joint. This study highlights that different RA stages corresponds to different gut-joint axes that may contribute to different symptoms of patients. The stage-specific intervention in microbial dysbiosis and metabolic disorders was warranted for prognosis and prevention of RA progression.

Methods

The cohort for the metagenomic and metabolomic study

A total of 122 fecal and 122 serum samples were collected from 122 outpatients of the Shandong Provincial Qianfoshan Hospital (Jinan, Shandong, China). These

outpatients included 27 healthy individuals, 19 patients with osteoarthritis (OA), and 76 patients with rheumatoid arthritis (RA). RA patients 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[15], where RAS1, RAS2, RAS3, and RAS4 has a score of 6-7, 8, 9, and 10, respectively (**Supplementary Tables 1**). The score was evaluated by the sum of four categories[15]: A) joint involvement that refers to any swollen or tender joint on examination; B) Serology determined by values of rheumatoid factor and anti-citrullinated protein antibody; C) Acute-phase reactants determined by values of C-reactive protein and erythrocyte sedimentation rate; and D) Duration of symptoms. Subsequently, fecal samples were sequenced and serum samples were used to test the serum metabolites and inflammatory cytokine levels. Serum metabolites were tested using UHPLC-MS/MS, while inflammatory cytokines were quantified using the MESO SCALE DISCOVERY (MSD) Quick Plex S600MM multiplex assay. Cytokine levels of healthy individuals were not available because of their extremely low levels. 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 had no malignant tumor, no other rheumatic diseases such as ankylosing spondylitis, psoriasis, gout, no gastrointestinal diseases such as diarrhea, constipation, and hematochezia in the recent one month, no infections, no other comorbidity such as diabetes and hepatitis B.

The cohort for detection of bacteria in the joint synovial fluid

A total of 271 patients with RA of four distinct stages were recruited, including 52 RAS1, 66 RAS2, 67 RAS3, and 86 RAS4. RA patients were grouped into four RA stages according to the rheumatoid diagnostic score[15]. The patients met the aforementioned standard. Synovial fluid (SF) samples were collected aseptically from knee joints during therapeutic aspiration. SF samples were deposited in sterile tubes on ice and homogenized within five minutes of collection. A tube filled with sterile phosphate-buffered saline (PBS) was left open throughout the procedure and

subsequently processed in parallel with the samples as a negative control. The entire experiment was conducted in a completely sterile atmosphere. Each sample was immediately frozen and kept without heparin or hyaluronidase at –135 °C. For each patient, a total of 7ml SF was collected, of which 5ml was utilized for 16S rRNA gene sequencing, 1ml was used for bacteria isolation, and 1ml SF was prepared for scanning electron microscopy.

Metagenome sequencing and processing

Whole-genome shotgun sequencing of fecal samples was carried out on the Illumina Hiseq X Ten. All samples were paired-end sequenced with a 150-bp read length. Reads containing more than 40 base pairs with low-quality values of 38 or less were discarded. Reads containing more than 10 unidentified base pairs were discarded. Reads containing more than 15 base pairs that overlapped adapter sequences were discarded. Next, reads were mapped against the human genome using Bowtie2 (version 2.2.9). Those that were mapped were considered to be derived from the human genome and were discarded. Finally, unpaired reads were discarded. High-quality reads were used for the following analyses. MetaPhlAn2[18] was then used to generate microbial species profiles.

Contig assembly and ORF prediction

The reads were assembled into contigs using MEGAHIT[17] (version 1.2.6) with the minimum contig length set at 500 bp and default parameters for metagenomics. The open reading frames (ORFs) were predicted from the assembled contigs using Prodigal[50] (version 2.6.3) with default parameters. The ORFs of <100 bp were removed.

Non-redundant gene catalogue construction

The ORFs were clustered to remove redundancy using Cd-hit[51] (version 4.6.6) with a sequence identity threshold set at 0.95 and the alignment coverage set at 0.9, which

resulted in a catalogue of 4,047,645 nonredundant genes. The non-redundant genes were then collapsed into metagenomic species (MGS)[16,52], and grouped into KEGG functional modules[16], as described below.

Identification of metagenomic species

High-quality reads were mapped to the catalogue of nonredundant genes using Bowtie 2[53] (version 2.2.9) with default parameters. The abundance profile for each catalogue gene was calculated as the sum of uniquely mapped sequence reads, using 19M sequence reads per sample (downsized). The co-abundance clustering of the 4,047,645 genes was performed using canopy algorithm[52] (http://git.dworzynski.eu/mgs-canopy-algorithm), and 553 gene clusters that met the previously described criteria[52] and contained more than 700 genes were referred to as MGS. MGS that were present in at least 4 samples were used for the following analysis. The abundance profiles of MGS were determined as the medium gene abundance throughout the samples. MGS were taxonomically annotated by summing up the taxonomical annotation of their genes as described by Nielsen *et al*[52]. Each MGS gene was annotated by sequence similarity to the bacterial complete genome in NCBI Reference Sequence Database (BLASTN, E-value < 0.001).

Annotation of KEGG modules.

The catalogue of the nonredundant genes was functionally annotated to the KEGG database (release 94.0) by KofamKOALA (version 1.3.0)[54,55]. The produced KEGG Orthologs (KOs) were mapped to the KEGG modules annotation downloaded on 1 August 2020 from the KEGG BRITE database. KOs that were present in at least 4 samples were used for the following analysis. The KO abundance profile was calculated by summing the abundance of genes that were annotated to each of the KOs.

Leave-one-out analysis

Leave-one-out analysis was used to test which MGS was driving the observed associations between KEGG modules and arthritis. The calculation of the KO abundance was iterated excluding the genes from a different MGS, in each iteration. The effect of a given MGS on a specified association was defined as the change in median Spearman correlation coefficient between KOs and arthritis when genes from the respective MGS were left out, as previously described[16,56].

Multi-omics biomarkers

A random forest algorithm was performed using the R package “randomForest” on each of the three profiles independently, including 232 microbiome species, 6,224 KOs, and 277 metabolites. 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 fid models over different tuning parameters to determine the “mtry” for random forest algorithm.

Correlation of phenotypes and KEGG modules

The clinical phenotypes including types of arthritis (Healthy = 0, OA = 1, RA = 2) were used in the correlation analysis with KEGG modules. The associations of clinical phenotypes with KEGG modules were determined by if Spearman correlations of the phenotype with the abundances of KOs in the given KEGG module were significantly higher or lower (Mann–Whitney U-test FDR < 0.1) than with the abundances of all other KOs. The phenotypes adjusted by age and gender were also tested using partial Spearman correlation. The union set of the significant associations between KEGG modules and phenotypes/phenotypes adjusted by age and gender were used for the following analysis.

Gavage experiments using ascorbate in the CIA mouse model

Nine healthy seven-week-old DBA/1 mice weighing 20g were fed in an ultra-clean animal laboratory (SPF grade) with a humidity of 55% and a temperature of 26°C. CIA models were constructed and established as described before[57]. Nine mice were then divided into three groups (three mice per group), including normal DBA/1 mice and two groups of DBA/1 mice with CIA. Three-month gavage (0.3ml/d) to three groups of mice was conducted, including 1) 0.9% normal saline to normal DBA/1 mice, 2) 0.9% normal saline to DBA/1 mice with CIA, and 3) 100ng/ul ascorbate to DBA/1 mice with CIA. After three-month gavage, the mice serum TNF-α level and the IL-6 level were tested using ELISA kit (mlbio, China). Mice were then killed and preserved in 4% formalin for two days. Micro-CT (QuantumGX, PerkinElmer, UnitedStates) was used to perform scanning and three-dimensional structural reconstruction of the joints. The settings were set to 209m, 90kV X-ray tube voltage, 160uA of the current, and 3 minutes of the scan time. The angle of the X-ray scan rotated 180 degrees.

16S rRNA gene sequencing and processing

Bacterial DNA was extracted from 271 5ml synovial fluid samples. The tube containing PBS serves as environmental control. 86 SF samples from patients in RAS4 had bacteria DNA content (≥10ng) (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[58]. Sequence analysis was performed using the Quantitative Insights into Microbial Ecology (QIIME, version 1.9.1) software[59]. Chimeric sequences were removed using de novo chimera detection of USEARCH 6.1[60]. Sequences were clustered against the 2013 Greengenes (13_8 release) ribosomal database's 97% reference data set. Sequences that did not match any entries in this reference were subsequently clustered into de novo OTUs at 97% similarity

with UCLUST[60]. Taxonomy was assigned to all OTUs using the RDP classifier[61] against Greengenes reference data set.

Isolation of bacteria in the joint synovial fluid

Due to that pH of joint SF was 7.6, we used Luria-Bertani (LB) broth medium with a pH of 7.6 for cultivation. The procedures were as follows: 1) sterilized LB medium (pH7.6) and inoculation equipment were placed in an anaerobic glove box (Ruskinn Concept 400) for deaeration one day in advance; 2) 1ml SF samples per stage of RA were used for bacterial cultivation. 3) The SF sample was serially diluted with sterilized water, plated onto LB (0.5% yeast extract, 1% tryptone, 1% sodium chloride) agar medium, and then incubated at 37°C for 72 h to obtain single colonies. Then, three colonies per plate were randomly selected and streaked three consecutive times on LB agar medium to obtain a pure culture, which was named isolate SF1 to SF9, respectively. To identify isolate SF1 to SF9, a partial fragment of 16S rDNA was amplified with the primer pair 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3'), and then DNA sequencing was performed for preliminary identification.

Scanning electron microscopy for bacteria in the joint synovial fluid

The SF sample was filtered through the membrane (special membrane for flow cytometry) to remove solids and large particles (Filtrate A). Filtrate A was then filtered through a 0.45 um membrane to remove most human cells (Filtrate B). Filtrate B was then filtered through a 0.22 um filter membrane to enrich bacteria and subsequently washed with sterilized ultra-pure water to obtain Filtrate C. Filtrate C was soaked in 2.5% glutaraldehyde for 4 hours at room temperature. Filtrate C was then washed three times with 0.1 M PBS buffer and treated with 1% osmic acid for 4h. Filtrate C was then dehydrated in ethanol, vacuum dried by tert-butyl alcohol, coated with gold, and imaged with a scanning electron microscope (ZEISS Sigma 300).

Statistical analyses

The correlations between vectors such as microbial species, metabolites, and the levels of cytokines, were tested using Spearman correlation or partial Spearman correlation with age and gender adjusted (*P* values), corrected by FDR (q values). For categorical metadata, samples were pooled into bins (RA/OA/Healthy, RAS1/RAS2/RAS3/RAS4) and significant features were identified using Mann-Whitney-Wilcoxon Test (*P* values) with Benjamini and Hochberg correction (FDR, q values).

Data availability

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).

Acknowledgements

This work was partially supported by the National Natural Science Foundation of China (grant numbers 31871334, 82003766, 32071465, and 31671374) and the Academic Promotion Project of Shandong First Medical University (grant number 2019LJ001).

Author contributions

M.C., Y.Z., J.H., and K.N. designed the study. M.C., Y.Z., Y.C., Y. Zha, S.L., G.C., and M.L. collected samples and conducted experiments. M.C., Y.Z conducted data analysis and produced the figures and tables. M.C., Y.Z., J.H., and K.N. wrote the manuscript. All authors revised the manuscript. M.C., Y.Z., J.H., and K.N. supervised the study.

Ethics statement

The study was approved by the Ethics Committee of The First Affiliated Hospital of Shandong First Medical University (NO.2017-02).

Competing interests

The authors declare that they have no competing interests

References

1. Almutairi, K., Nossent, J., Preen, D., Keen, H. & Inderjeeth, C. The global prevalence of rheumatoid arthritis: a meta-analysis based on a systematic review. *Rheumatol. Int.* **41**, 863–877 (2021).
2. Barhum, L. What Does Rheumatoid Arthritis Progression Look Like? https://www.verywellhealth.com/rheumatoid-arthritis-stages-of-progression-4768891 (2021).
3. Smolen, J.S., Aletaha, D. & McInnes, I.B. Rheumatoid arthritis. *Lancet* **388**, 2023–2038 (2016).
4. Wells, P.M., *et al.* Associations between gut microbiota and genetic risk for rheumatoid arthritis in the absence of disease: a cross-sectional study. *Lancet Rheumatol* **2**, e418–e427 (2020).
5. Zaiss, M.M., Joyce Wu, H.J., Mauro, D., Schett, G. & Ciccia, F. The gut-joint axis in rheumatoid arthritis. *Nat. Rev. Rheumatol.* **17**, 224–237 (2021).
6. Jubair, W.K., *et al.* Modulation of Inflammatory Arthritis in Mice by Gut Microbiota Through Mucosal Inflammation and Autoantibody Generation. *Arthritis Rheumatol* **70**, 1220–1233 (2018).
7. Takahashi, D., *et al.* Microbiota-derived butyrate limits the autoimmune response by promoting the differentiation of follicular regulatory T cells. *EBioMedicine* **58**, 102913 (2020).
8. Smith, P.M., *et al.* The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* **341**, 569–573 (2013).
9. Zhao, Y., *et al.* Hidden link in gut-joint axis: gut microbes promote rheumatoid arthritis at early stage by enhancing ascorbate degradation. *Gut* (2021).
10. Pianta, A., *et al.* Evidence of the Immune Relevance of Prevotella copri, a Gut Microbe, in Patients With Rheumatoid Arthritis. *Arthritis Rheumatol* **69**, 964–975 (2017).
11. Chen, J., *et al.* An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis. *Genome Med.* **8**, 43 (2016).
12. Round, J.L. & Mazmanian, S.K. Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 12204–12209 (2010).
13. Chun, E., *et al.* Metabolite-Sensing Receptor Ffar2 Regulates Colonic Group 3 Innate Lymphoid Cells and Gut Immunity. *Immunity* **51**, 871–884.e876 (2019).
14. Maeda, Y. & Takeda, K. Role of Gut Microbiota in Rheumatoid Arthritis. *J Clin Med* **6**, 60

(2017).

15. Aletaha, D., *et al.* 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheumatol* **62**, 2569–2581 (2010).

16. Pedersen, H.K., *et al.* A computational framework to integrate high-throughput '-omics' datasets for the identification of potential mechanistic links. *Nat Protoc* **13**, 2781-2800 (2018).

17. Li, D., *et al.* MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* **102**, 3–11 (2016).

18. Truong, D.T., *et al.* MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat. Methods* **12**, 902–903 (2015).

19. de Oliveira, G.L.V., Leite, A.Z., Higuchi, B.S., Gonzaga, M.I. & Mariano, V.S. Intestinal dysbiosis and probiotic applications in autoimmune diseases. *Immunology* **152**, 1–12 (2017).

20. Lugli, G.A., *et al.* Decoding the Genomic Variability among Members of the Bifidobacterium dentium Species. *Microorganisms* **8**, 1720 (2020).

21. Silvestre-Rangil, J., Bagán, L., Silvestre, F.J. & Bagán, J.V. Oral manifestations of rheumatoid arthritis. A cross-sectional study of 73 patients. *Clin Oral Investig* **20**, 2575-2580 (2016).

22. Marriott, D., Stark, D. & Harkness, J. Veillonella parvula discitis and secondary bacteremia: a rare infection complicating endoscopy and colonoscopy? *J. Clin. Microbiol.* **45**, 672–674 (2007).

23. Balakrishnan, B., Luckey, D. & Taneja, V. Autoimmunity-Associated Gut Commensals Modulate Gut Permeability and Immunity in Humanized Mice. *Mil. Med.* **184**, 529–536 (2019).

24. Tito, R.Y., *et al.* Brief Report: Dialister as a Microbial Marker of Disease Activity in Spondyloarthritis. *Arthritis Rheumatol* **69**, 114–121 (2017).

25. Severmann, A.C., *et al.* An altered heparan sulfate structure in the articular cartilage protects against osteoarthritis. *Osteoarthr Carti* **28**, 977–987 (2020).

26. Henrotin, Y., Mathy, M., Sanchez, C. & Lambert, C. Chondroitin sulfate in the treatment of osteoarthritis: from in vitro studies to clinical recommendations. *Ther. Adv. Musculoskelet. Dis.* **2**, 335–348 (2010).

27. Jura-Półtorak, A., *et al.* Alterations of plasma glycosaminoglycan profile in patients with rheumatoid arthritis in relation to disease activity. *Clin. Chim. Acta* **433**, 20–27 (2014).

28. Shamdani, S., *et al.* Heparan sulfate functions are altered in the osteoarthritic cartilage. *Arthritis Res. Ther.* **22**, 283 (2020).

29. Pattison, D.J., *et al.* Vitamin C and the risk of developing inflammatory polyarthritis: prospective nested case-control study. *Ann. Rheum. Dis.* **63**, 843-847 (2004).

30. Carr, A.C. & Maggini, S. Vitamin C and Immune Function. *Nutrients* **9**(2017).

31. Abrams, E. & Sandson, J. EFFECT OF ASCORBIC ACID ON RHEUMATOID SYNOVIAL FLUID. *Ann. Rheum. Dis.* **23**, 295–299 (1964).

32. Abdelkarem, H.M., Fadda, L.H., El-Sayed, E.M. & Radwan, O.K. Potential Role of L-Arginine and Vitamin E Against Bone Loss Induced by Nano-Zinc Oxide in Rats. *J. Diet. Suppl.* **15**, 300–310 (2018).

33. Fiore, C.E., *et al.* L-arginine prevents bone loss and bone collagen breakdown in cyclosporin A-treated rats. *Eur. J. Pharmacol.* **408**, 323–326 (2000).

34. Manfredo Vieira, S., *et al.* Translocation of a gut pathobiont drives autoimmunity in mice and

humans. *Science* **359**, 1156–1161 (2018).

35. Mårdh, P.A., Nilsson, F.J. & Bjelle, A. Mycoplasmas and bacteria in synovial fluid from patients with arthritis. *Ann. Rheum. Dis.* **32**, 319–325 (1973).

36. Zhao, Y., *et al.* Detection and characterization of bacterial nucleic acids in culture-negative synovial tissue and fluid samples from rheumatoid arthritis or osteoarthritis patients. *Sci. Rep.* **8**, 14305 (2018).

37. Siala, M., *et al.* Analysis of bacterial DNA in synovial tissue of Tunisian patients with reactive and undifferentiated arthritis by broad-range PCR, cloning and sequencing. *Arthritis Res. Ther.* **10**, R40 (2008).

38. Cox, C.J., Kempsell, K.E. & Gaston, J.S. Investigation of infectious agents associated with arthritis by reverse transcription PCR of bacterial rRNA. *Arthritis Res. Ther.* **5**, R1–8 (2003).

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

40. Pucino, V., *et al.* Lactate Buildup at the Site of Chronic Inflammation Promotes Disease by Inducing CD4(+) T Cell Metabolic Rewiring. *Cell Metab.* **30**, 1055–1074 (2019).

41. Rogatzki, M.J., Ferguson, B.S., Goodwin, M.L. & Gladden, L.B. Lactate is always the end product of glycolysis. *Front. Neurosci.* **9**, 22 (2015).

42. Brown, N.A., Holt, D. & Webb, M. The teratogenicity of methoxyacetic acid in the rat. *Toxicol. Lett.* **22**, 93–100 (1984).

43. Miller, R.R., Carreon, R.E., Young, J.T. & McKenna, M.J. Toxicity of methoxyacetic acid in rats. *Fundam Appl Toxicol* **2**, 158–160 (1982).

44. Sparks, N.R. *The Embryotoxic Effects of Harm Reduction Tobacco Products on Osteoblasts Developing from Human Embryonic Stem Cellss*, (University of California, Riverside, 2018).

45. Li, P., *et al.* Metabolic Alterations in Older Women With Low Bone Mineral Density Supplemented With Lactobacillus reuteri. *JBMR Plus* **5**, e10478 (2021).

46. Picchianti-Diamanti, A., *et al.* Analysis of Gut Microbiota in Rheumatoid Arthritis Patients: Disease-Related Dysbiosis and Modifications Induced by Etanercept. *Int. J. Mol. Sci.* **19**, 2938 (2018).

47. Schirmer, M., *et al.* Linking the Human Gut Microbiome to Inflammatory Cytokine Production Capacity. *Cell* **167**, 1125–1136.e1128 (2016).

48. Kverka, M., *et al.* Oral administration of Parabacteroides distasonis antigens attenuates experimental murine colitis through modulation of immunity and microbiota composition. *Clin Exp Immunol* **163**, 250–259 (2011).

49. Koh, G.Y., Kane, A.V., Wu, X. & Crott, J.W. Parabacteroides distasonis attenuates tumorigenesis, modulates inflammatory markers and promotes intestinal barrier integrity in azoxymethane-treated A/J mice. *Carcinogenesis* **41**, 909–917 (2020).

50. Hyatt, D., *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform* **11**, 119 (2010).

51. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).

52. Nielsen, H.B., *et al.* Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat. Biotechnol.* **32**, 822–828 (2014).

53. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

54. Kanehisa, M., *et al*. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res*. **42**, D199–205 (2014).
55. Aramaki, T., *et al.* KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **36**, 2251–2252 (2020).
56. Pedersen, H.K., *et al*. Human gut microbes impact host serum metabolome and insulin sensitivity. *Nature* **535**, 376–381 (2016).
57. Song, G., *et al*. CD109 regulates the inflammatory response and is required for the pathogenesis of rheumatoid arthritis. *Ann. Rheum. Dis*. **78**, 1632–1641 (2019).
58. Magoč, T. & Salzberg, S.L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963 (2011).
59. Caporaso, J.G., *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
60. Edgar, R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
61. Cole, J.R., *et al.* The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res*. **37**, D141–145 (2009).

Figure legends

Fig. 1 | The analysis pipeline for detecting stage-specific gut-arthritis axis. Gut microbiome whole-genome sequencing (WGS) data and serum metabolomic data from 27 healthy controls, 76 rheumatoid arthritis (RA), and 19 osteoarthritis (OA) patients were used in the analysis. RA samples were divided into four stages including RAS1 (n = 15), RAS2 (n = 21), RAS3 (n = 18), RAS4 (n = 22) according to the rheumatoid diagnostic score. **a**, Microbial genes, predicted from WGS sequences, were clustered into metagenomic species (MGS) for taxonomic annotation and mapped to KEGG orthology for functional annotation. Ten KEGG modules that were significantly correlated with arthritis (Healthy = 0, OA = 1, RA = 2, $P < 0.05$, FDR < 0.05, Mann-Whitney-Wilcoxon test) were remained to describe the stage-specific functional pattern. Additionally, the leave-one-out analysis was used to determine the MGS that most drove the correlation of these ten KEGG modules with arthritis (Methods). KEGG maps of these specific modules, their KOs' abundances, and the corresponding driving MGS abundances were plotted to show the stage-specific functional driving genes variation. **b**, A total of 232 classified microbial species, identified from WGS sequences, were used to describe the stage-specific taxonomic

pattern. **c**, In total, 277 serum metabolites were used to describe the stage-specific metabolomic pattern. **d**, Random forest algorithm was performed on 6224 KOs, 232 microbial species, and 277 serum metabolites from (a–c) to detect the stage-specific biomarkers at the multi-omics level and the AUROC in classifying stage-specific arthritis samples from healthy controls.

Fig. 2 | Distinct stage-specific taxonomic patterns in arthritis progression. a, Linear discriminant analysis to separate the samples of healthy, OA, and RA groups. The length and direction of the arrows indicate the normalized scalings for each of the microbial species. **b**, Each row represents a species with significantly different abundances ($P < 0.05$, FDR < 0.05) between the healthy and RA groups or between the healthy and OA groups. Each column represents a sample from one of the groups including the healthy (n = 27), RAS1 (n = 15), RAS2 (n = 21), RAS3 (n = 18), RAS4 (n = 22), and OA (n = 19) groups. The color of each heatmap unit represents the scaled abundance of a certain species in a specific sample. Species are colored for significant elevation (red) or depletion (green) in the arthritis groups, in comparison with the healthy groups. **c**, Three apices of the triangle in the ternary diagram represent components of healthy, OA, and RA groups. Microbial species in (b) are plotted as dots inside the triangle, the size of which indicates the relative abundance. For each dot, moving away from the apex, the proportion of the component at the apex decreases as the proportion of the remaining two increases. **d**, In total, 232 classified species whose relative abundance was more than 1×10^{-6} are shown in a phylogenetic tree, grouped in the phyla Fusobacteria, Proteobacteria, Actinobacteria, Bacteroidetes, and Firmicutes. In the outer circles, species at each arthritis stage are marked for significant elevation ($P < 0.05$, orange; FDR < 0.05, red) or depletion ($P < 0.05$, light green; FDR < 0.05, dark green) in abundance, as compared to those of healthy controls. The innermost circle shows species relative abundances averaged over all samples. Box plots show the relative abundances of species recorded in (b) and (c), and others that have been correlated with RA in previous work. The order and color of the box are consistent with the stage circle. Significant changes (elevation

and depletion) 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 $P < 0.05$. Boxes represent the interquartile range between the first and third quartiles and the line inside represents the median. Whiskers denote the lowest and highest values within the 1.5 × interquartile range from the first and third quartiles, respectively. All the statistical significance was calculated by the Mann-Whitney-Wilcoxon test. Microbial species shared in (b−d) are marked with consistent circular marks.

Fig. 3 | Distinct stage-specific functional patterns in arthritis progression. 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 to the healthy controls (n = 27). **a**, Relative abundance of KO genes in the KEGG modules that were significantly correlated with arthritis ($P < 0.05$, FDR < 0.05). KO genes with a prevalence of 5% or higher are shown. Significant changes (elevation and depletion) 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 $P < 0.05$. **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 to healthy controls. 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 colored according to the values. Each KO gene is composed of MGS genes represented by circles. The sizes and colors 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. Significant changes (elevation and depletion) 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

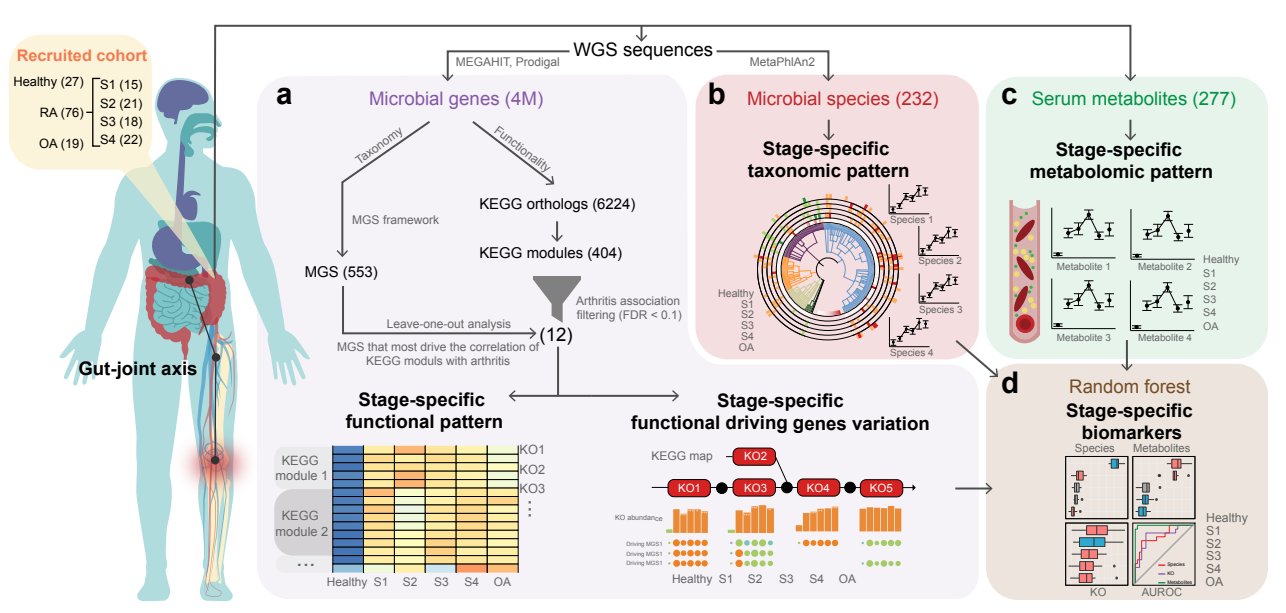
$P < 0.01$; −, depletion at $P < 0.05$. The boxes represent 25th–75th percentiles, black lines indicate the median, and whiskers extend to the maximum and minimum values within 1.5 × the interquartile range. All the statistical significance was calculated by the Mann-Whitney-Wilcoxon test.

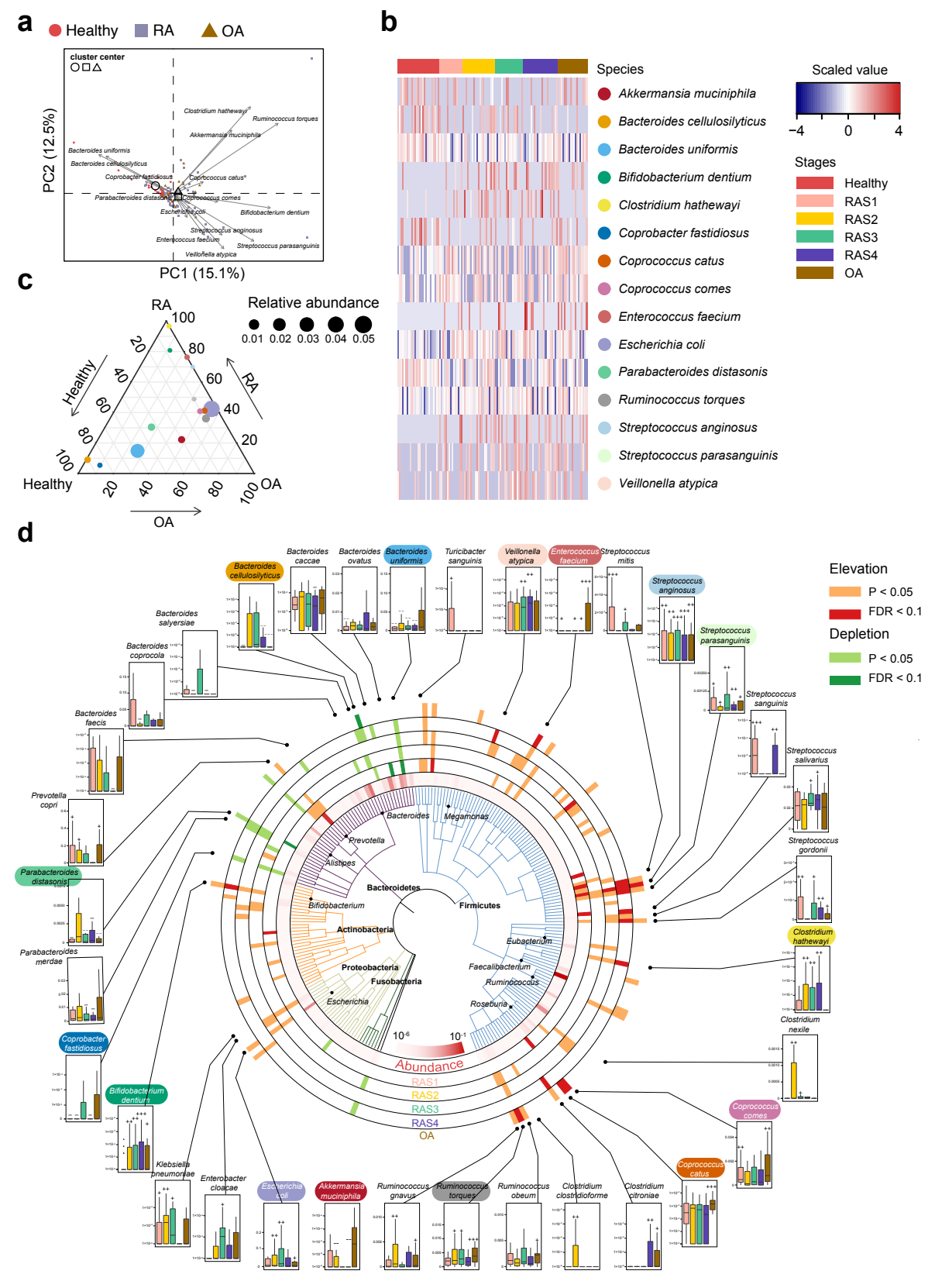
Fig. 4 | Stage-specific biomarkers for discriminating arthritis progression from healthy status at a multi-omics level. A random forest algorithm was performed on 6224 KOs, 232 microbial species, and 277 serum metabolites to detect the stage-specific biomarkers in the arthritis progression including RAS1 (**a**), RAS2 (**b**), RAS3 (**c**), RAS4 (**d**), and OA (**e**). In each of the panels, the Gini importance of the top five most discriminant biomarkers of microbial species, KOs, or metabolites are displayed. Boxes represent the interquartile range between the first and third quartiles and the line inside represents the median. Whiskers denote the lowest and highest values within the 1.5 × interquartile range from the first and third quartiles, respectively. Boxes are marked in a specific color 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 biomarkers in each of the arthritis stages compared to the healthy group. The ROC curves of the random-forest model using microbial species, KOs, or metabolites were plotted, with AUC calculated by 10 randomized 10-fold cross-validation. The color of the curve represents the categories of the used features.

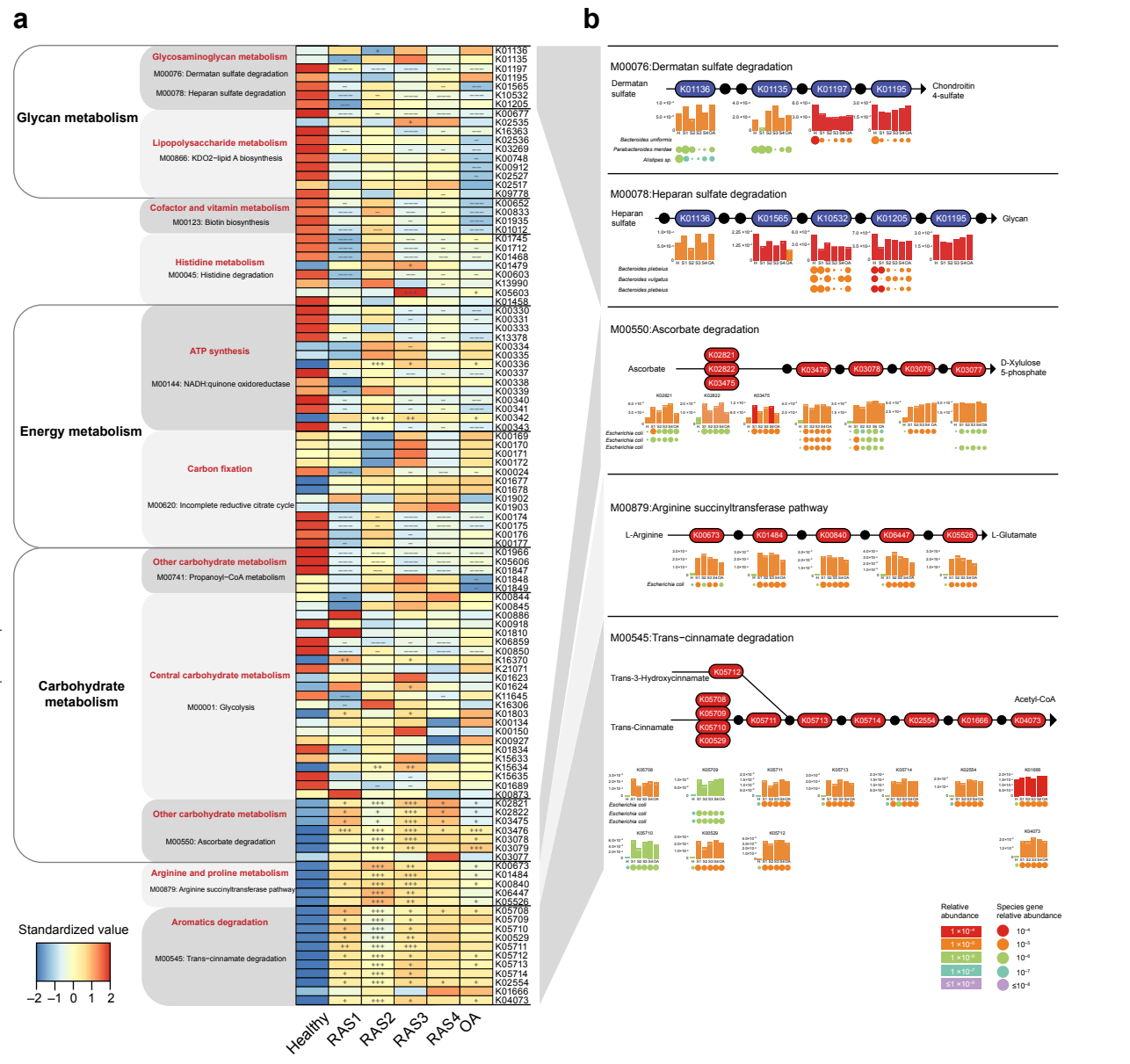
Fig. 5 | Stage-specific gut microbial species, serum cytokines, and serum metabolites alteration in gut-joint axis. a, The alteration of gut microbiome taxonomic and functional composition would interact with human joint through serum metabolites perturbation. The most dominant elevation and depletion of microbial functionality in arthritis, or microbial species or serum metabolites at each arthritis stage are shown, as compared to those in the healthy group. The perturbated metabolites and elevated serum cytokines such as IL-6 and TNF-α were transited to human joints to promote inflammation and bone erosion. Different arthritis stages had inflammation or bone erosion of different severity and corresponding different

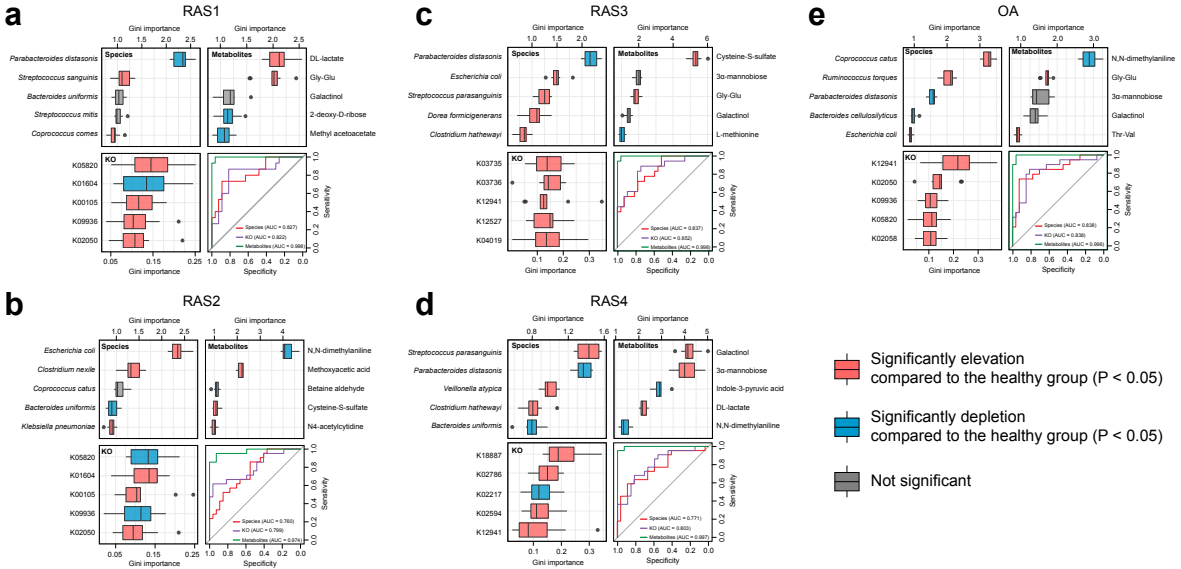
microbial patterns. **b**, The dot plots show stage-specific abundance or concentration (mean ± s.e.m.) of microbial species, serum cytokines, and serum metabolites, which are specified in (a). Four RA stages are connected to display the variance. For microbial species and serum metabolites, dots are colored differently if the features are significantly elevated (red) or significantly depleted (blue), as compared to those of the healthy group. For serum cytokines, dots of RA stages are colored in red for significant elevation compared to the OA group. ***, $P < 0.005$; **, $P < 0.01$; *, $P < 0.05$, Mann-Whitney-Wilcoxon test.

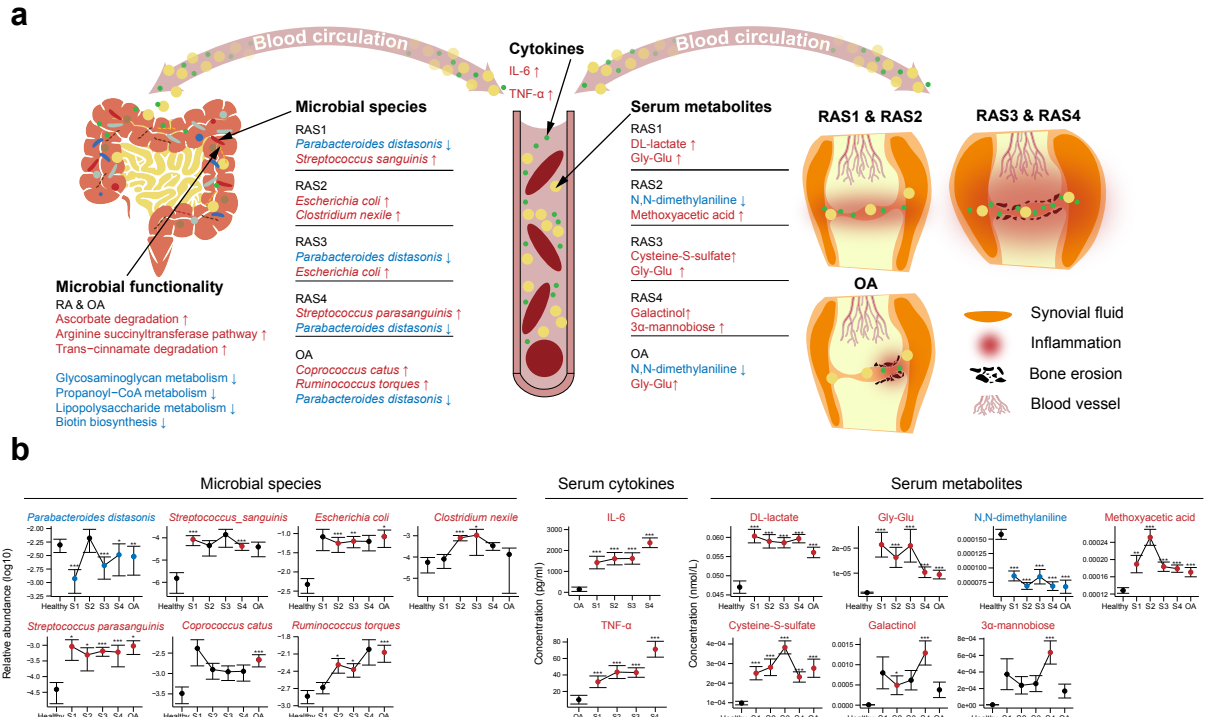
Fig. 6 | The RA-stage-specific changes were related to RA pathogenesis. a, Certain potential mechanisms by which gut microbial dysbiosis is involved in RA pathogenesis are displayed. Gut microbial variation drove the changes in microbial functionality that influenced the metabolites, which might promote RA progression via hurting hone tissue and increasing inflammation. The driving species, functions, and metabolites were extracted from **Fig. 3b**. The red or blue box of metabolites represents their speculated elevation or depletion according to the functional variation. The dotted line represents the speculated effects of microbial and metabolic variation on arthritis pathogenesis, with a bracket including the literature. **b**, The stage-specific changes at each of the RA stages that would promote RA progression. The major contributions of each RA stage, reflected by multi-omics, are displayed. The grey box showed that although these processes were the most notable changes at each of the RA stages, they kept ongoing throughout the RA progression. (a) characterized a few links between microbial dysbiosis and the joint of (b) throughout the RA progression.

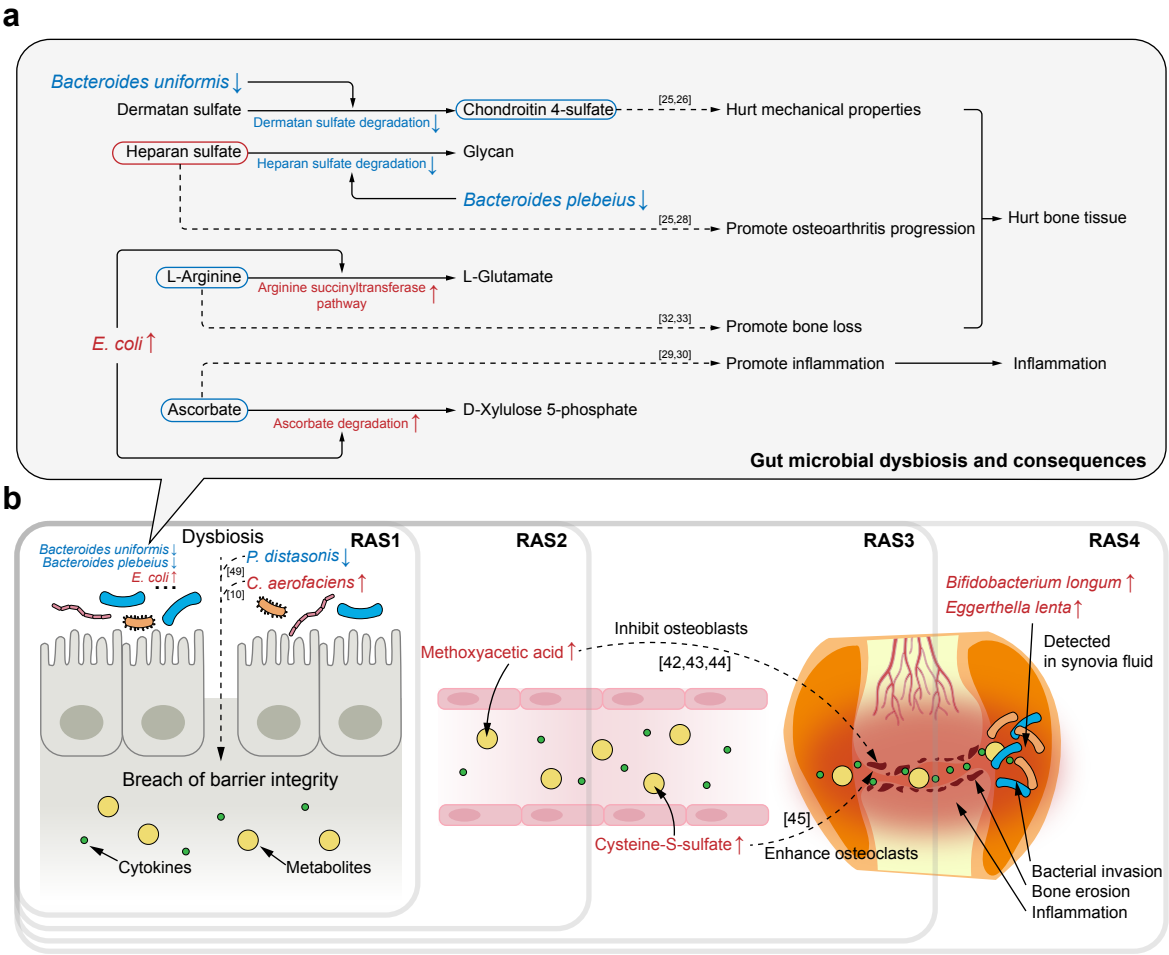

Recruited cohort
Healthy (27)
RA (76)
S1 (15)
S2 (21)
S3 (18)
S4 (22)
OA (19)
Gut-joint axis
WGS sequences
MEGAHIT, Prodigal
MetaPhlAn2
a
Microbial genes (4M)
Taxonomy
Functionality
MGS framework
MGS (553)
KEGG orthologs (6224)
KEGG modules (404)
Arthritis association filtering (FDR < 0.1)
Leave-one-out analysis
MGS that most drive the correlation of KEGG moduls with arthritis
(12)
Stage-specific functional pattern
KEGG module 1
KEGG module 2
KO1
KO2
KO3
Healthy S1 S2 S3 S4 OA
Stage-specific functional driving genes variation
KEGG map
KO1
KO2
KO3
KO4
KO5
KO abundance
Driving MGS1
Healthy S1 S2 S3 S4 OA
b
Microbial species (232)
Stage-specific taxonomic pattern
Species 1
Species 2
Species 3
Species 4
Healthy
S1
S2
S3
S4
OA
c
Serum metabolites (277)
Stage-specific metabolomic pattern
Metabolite 1
Metabolite 2
Metabolite 3
Metabolite 4
Healthy
S1
S2
S3
S4
OA
d
Random forest
Stage-specific biomarkers
Species
Metabolites
KO
AUROC
Species
KO
Metabolites
Healthy
S1
S2
S3
S4
OA


a
Healthy
RA
OA
cluster center
PC2 (12.5%)
PC1 (15.1%)
Clostridium hathewayi
Ruminococcus torques
Akkermansia muciniphila
Bacteroides uniformis
Bacteroides cellulosilyticus
Coprobacter fastidiosus
Coprococcus catus*
Parabacteroides distasonis
Coprococcus comes
Escherichia coli
Bifidobacterium dentium
Streptococcus anginosus
Enterococcus faecium
Streptococcus parasanguinis
Veillonella atypica
b
Species
Akkermansia muciniphila
Bacteroides cellulosilyticus
Bacteroides uniformis
Bifidobacterium dentium
Clostridium hathewayi
Coprobacter fastidiosus
Coprococcus catus
Coprococcus comes
Enterococcus faecium
Escherichia coli
Parabacteroides distasonis
Ruminococcus torques
Streptococcus anginosus
Streptococcus parasanguinis
Veillonella atypica
Scaled value
−4
0
4
Stages
Healthy
RAS1
RAS2
RAS3
RAS4
OA
c
RA
Healthy
OA
Relative abundance
0.01
0.02
0.03
0.04
0.05
d
Elevation
P < 0.05
FDR < 0.1
Depletion
P < 0.05
FDR < 0.1
Bacteroides cellulosilyticus
Bacteroides caccae
Bacteroides ovatus
Bacteroides uniformis
Turicibacter sanguinis
Veillonella atypica
Enterococcus faecium
Streptococcus mitis
Streptococcus anginosus
Streptococcus parasanguinis
Streptococcus sanguinis
Streptococcus salivarius
Streptococcus gordonii
Clostridium hathewayi
Clostridium nexile
Coprococcus comes
Coprococcus catus
Clostridium citroniae
Clostridium clostridioforme
Ruminococcus obeum
Ruminococcus torques
Ruminococcus gnavus
Akkermansia muciniphila
Escherichia coli
Enterobacter cloacae
Klebsiella pneumoniae
Bifidobacterium dentium
Coprobacter fastidiosus
Parabacteroides merdae
Parabacteroides distasonis
Prevotella copri
Bacteroides faecis
Bacteroides coprocola
Bacteroides salyersiae
Bacteroides
Megamonas
Prevotella
Alistipes
Bacteroidetes
Firmicutes
Bifidobacterium
Actinobacteria
Eubacterium
Proteobacteria
Faecalibacterium
Fusobacteria
Ruminococcus
Escherichia
Roseburia
10^{-6}
10^{-1}
Abundance
RAS1
RAS2
RAS3
RAS4
OA


a
Glycan metabolism
Glycosaminoglycan metabolism
M00076: Dermatan sulfate degradation
M00078: Heparan sulfate degradation
Lipopolysaccharide metabolism
M00866: KDO2−lipid A biosynthesis
Cofactor and vitamin metabolism
M00123: Biotin biosynthesis
Histidine metabolism
M00045: Histidine degradation
Energy metabolism
ATP synthesis
M00144: NADH;quinone oxidoreductase
Carbon fixation
M00620: Incomplete reductive citrate cycle
Carbohydrate metabolism
Other carbohydrate metabolism
M00741: Propanoyl−CoA metabolism
Central carbohydrate metabolism
M00001: Glycolysis
Other carbohydrate metabolism
M00550: Ascorbate degradation
Arginine and proline metabolism
M00879: Arginine succinyltransferase pathway
Aromatics degradation
M00545: Trans−cinnamate degradation
K01136
K01135
K01197
K01195
K00677
K02535
K16363
K02536
K03269
K00748
K00912
K02527
K02517
K09778
K00652
K00833
K01935
K01012
K01745
K01712
K01468
K01479
K00603
K13990
K05603
K01458
K00330
K00331
K00333
K13378
K00334
K00335
K00336
K00337
K00338
K00339
K00340
K00341
K00342
K00343
K00169
K00170
K00171
K00172
K00024
K01677
K01678
K01902
K01903
K00174
K00175
K00176
K00177
K01966
K05606
K01847
K01848
K01849
K00844
K00845
K00886
K00918
K01810
K06859
K06850
K16370
K21071
K01623
K01624
K11645
K16306
K01803
K00134
K00150
K00927
K01834
K15633
K15634
K15635
K01689
K00873
K02821
K02822
K03475
K03476
K03078
K03079
K03077
K00673
K01484
K00840
K06447
K05526
K05708
K05709
K05710
K00529
K05711
K05712
K05713
K05714
K02554
K01666
K04073
Standardized value
−2 −1 0 1 2
Healthy
RAS1
RAS2
RAS3
RAS4
OA
b
M00076:Dermatan sulfate degradation
Dermatan sulfate
Chondroitin 4-sulfate
Bacteroides uniformis
Parabacteroides merdae
Alistipes sp.
M00078:Heparan sulfate degradation
Heparan sulfate
Glycan
Bacteroides plebeius
Bacteroides vulgatus
Bacteroides plebeius
M00550:Ascorbate degradation
Ascorbate
D-Xylulose 5-phosphate
Escherichia coli
Escherichia coli
Escherichia coli
M00879:Arginine succinyltransferase pathway
L-Arginine
L-Glutamate
Escherichia coli
M00545:Trans−cinnamate degradation
Trans-3-Hydroxycinnamate
Trans-Cinnamate
Acetyl-CoA
Escherichia coli
Escherichia coli
Escherichia coli
Relative abundance
Species gene relative abundance


a
RAS1
Gini importance
Species
Parabacteroides distasonis
Streptococcus sanguinis
Bacteroides uniformis
Streptococcus mitis
Coprococcus comes
Metabolites
DL-lactate
Gly-Glu
Galactinol
2-deoxy-D-ribose
Methyl acetoacetate
KO
K05820
K01604
K00105
K09936
K02050
Species (AUC = 0.827)
KO (AUC = 0.820)
Metabolites (AUC = 0.968)
Sensitivity
Specificity
b
RAS2
Escherichia coli
Clostridium nexile
Coprococcus catus
Bacteroides uniformis
Klebsiella pneumoniae
N,N-dimethylaniline
Methoxyacetic acid
Betaine aldehyde
Cysteine-S-sulfate
N4-acetylcytidine
Species (AUC = 0.760)
KO (AUC = 0.799)
Metabolites (AUC = 0.974)
c
RAS3
Parabacteroides distasonis
Escherichia coli
Streptococcus parasanguinis
Dorea formicigenerans
Clostridium hathewayi
Cysteine-S-sulfate
3α-mannobiose
Gly-Glu
Galactinol
L-methionine
K03735
K03736
K12941
K12527
K04019
Species (AUC = 0.837)
KO (AUC = 0.832)
Metabolites (AUC = 0.966)
d
RAS4
Streptococcus parasanguinis
Parabacteroides distasonis
Veillonella atypica
Clostridium hathewayi
Bacteroides uniformis
Galactinol
3α-mannobiose
Indole-3-pyruvic acid
DL-lactate
N,N-dimethylaniline
K18887
K02786
K02217
K02594
K12941
Species (AUC = 0.771)
KO (AUC = 0.803)
Metabolites (AUC = 0.987)
e
OA
Coprococcus catus
Ruminococcus torques
Parabacteroides distasonis
Bacteroides cellulosilyticus
Escherichia coli
N,N-dimethylaniline
Gly-Glu
3α-mannobiose
Galactinol
Thr-Val
K12941
K02050
K09936
K05820
K02058
Species (AUC = 0.838)
KO (AUC = 0.838)
Metabolites (AUC = 0.995)
Significantly elevation compared to the healthy group (P < 0.05)
Significantly depletion compared to the healthy group (P < 0.05)
Not significant


a
Blood circulation
Blood circulation
Cytokines
IL-6 ↑
TNF-α ↑
Microbial species
RAS1
Parabacteroides distasonis ↓
Streptococcus sanguinis ↑
RAS2
Escherichia coli ↑
Clostridium nexile ↑
RAS3
Parabacteroides distasonis ↓
Escherichia coli ↑
RAS4
Streptococcus parasanguinis ↑
Parabacteroides distasonis ↓
OA
Coprococcus catus ↑
Ruminococcus torques ↑
Parabacteroides distasonis ↓
Serum metabolites
RAS1
DL-lactate ↑
Gly-Glu ↑
RAS2
N,N-dimethylaniline ↓
Methoxyacetic acid ↑
RAS3
Cysteine-S-sulfate↑
Gly-Glu ↑
RAS4
Galactinol↑
3α-mannobiose ↑
OA
N,N-dimethylaniline ↓
Gly-Glu↑
RAS1 & RAS2
RAS3 & RAS4
OA
Synovial fluid
Inflammation
Bone erosion
Blood vessel
Microbial functionality
RA & OA
Ascorbate degradation ↑
Arginine succinyltransferase pathway ↑
Trans−cinnamate degradation ↑
Glycosaminoglycan metabolism ↓
Propanoyl−CoA metabolism ↓
Lipopolysaccharide metabolism ↓
Biotin biosynthesis ↓
b
Microbial species
Serum cytokines
Serum metabolites
Parabacteroides distasonis
Streptococcus_sanguinis
Escherichia coli
Clostridium nexile
Streptococcus parasanguinis
Coprococcus catus
Ruminococcus torques
Relative abundance (log10)
Healthy S1 S2 S3 S4 OA
IL-6
TNF-α
Concentration (pg/ml)
OA S1 S2 S3 S4
DL-lactate
Gly-Glu
N,N-dimethylaniline
Methoxyacetic acid
Cysteine-S-sulfate
Galactinol
3α-mannobiose
Concentration (nmol/L)


a
Bacteroides uniformis ↓
Dermatan sulfate
Dermatan sulfate degradation ↓
Chondroitin 4-sulfate
[25,26]
Hurt mechanical properties
Heparan sulfate
Heparan sulfate degradation ↓
Glycan
Bacteroides plebeius ↓
[25,28]
Promote osteoarthritis progression
Hurt bone tissue
L-Arginine
Arginine succinyltransferase pathway ↑
L-Glutamate
[32,33]
Promote bone loss
E. coli ↑
[29,30]
Promote inflammation
Inflammation
Ascorbate
Ascorbate degradation ↑
D-Xylulose 5-phosphate
Gut microbial dysbiosis and consequences
b
Dysbiosis
RAS1
Bacteroides uniformis ↓
Bacteroides plebeius ↓
E. coli ↑
P. distasonis ↓
[49]
[10]
C. aerofaciens ↑
Breach of barrier integrity
Cytokines
Metabolites
RAS2
RAS3
RAS4
Inhibit osteoblasts
Methoxyacetic acid ↑
[42,43,44]
Cysteine-S-sulfate ↑
[45]
Enhance osteoclasts
Bifidobacterium longum ↑
Eggerthella lenta ↑
Detected in synovia fluid
Bacterial invasion
Bone erosion
Inflammation