

# 1 **Metagenomic Study Revealed the Potential Role of the Gut Microbiome in Gout**

2

3 Yongliang Chu<sup>1,6†</sup>, Yufen Huang<sup>2†</sup>, Qingchun Huang<sup>1†</sup>, Xuefeng Xie<sup>2†</sup>, Peng Wang<sup>2†</sup>, Junxia Li<sup>1</sup>,  
4 Lifeng Liang<sup>2</sup>, Xiaohong He<sup>1</sup>, Yiqi Jiang<sup>2,5</sup>, Maojie Wang<sup>1</sup>, Jianhua Yang<sup>2</sup>, Xiumin Chen<sup>1</sup>, Chu  
5 Zhou<sup>2</sup>, Yue Zhao<sup>1</sup>, Fen Ding<sup>2</sup>, Yi Zhang<sup>2</sup>, Xiaodong Wu<sup>1</sup>, Xueyuan Bai<sup>4</sup>, JiaQi Wu<sup>1</sup>, Xia Wei<sup>2</sup>,  
6 XiangHong Chen<sup>1</sup>, Xiaodong Fang<sup>1,2</sup>, Zhang Wang<sup>7</sup>, Qiang Gao<sup>1,2\*</sup>, Silong Sun<sup>1,2\*</sup>, and Runyue  
7 Huang<sup>1,3\*</sup>

8

- 9 1. The Second Affiliated Hospital of Guangzhou University of Chinese Medicine (Guangdong  
10 Provincial Hospital of Chinese Medicine), Guangzhou 510120 or 51006, China
- 11 2. BGI Genomics, BGI-Shenzhen, Shenzhen 518083, China
- 12 3. Guangdong Provincial Key Laboratory of Clinical Research on Traditional Chinese Medicine  
13 Syndrome, Guangzhou 510120, China
- 14 4. School of Environmental Science and Engineering, Sun Yat-Sen University, 135 Xingang West  
15 Street, Guangzhou 510275, China
- 16 5. Department of Computer Science, City University of Hong Kong, Tat Chee Avenue, Hong Kong,  
17 China
- 18 6. Zhuhai Branch, Guangdong Provincial Hospital of Chinese Medicine, Zhuhai, China
- 19 7. Institute of Ecological Science, School of Life Science, South China Normal University,  
20 Guangzhou, Guangdong Province, China, 510631

21

22 †These authors contributed equally to this work.

23 \*Corresponding authors:

24 Runyue Huang, MD, PhD

25 Section Rheumatology Research, The Second Affiliated Hospital of Guangzhou University of  
26 Chinese Medicine (Guangdong Provincial Hospital of Chinese Medicine), Guangzhou 510006,

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

1 China;

2 E-mail: ryhuang@gzucm.edu.cn

3

4 Silong Sun, PhD

5 BGI Genomics, BGI-Shenzhen, Shenzhen, 518083. China

6 E-mail: sunsilong@bgi.com

7

8 Qiang Gao, PhD

9 BGI Genomics, BGI-Shenzhen, Shenzhen, 518083. China

10 E-mail: gaoqiang@bgi.com

11

## 12 **ABSTRACT**

13 Emerging evidence has indicated an association between the gut microbiome and arthritis diseases  
14 including gout. This metagenomic study aims to investigate the possible role of gut microbiota in  
15 the development of gout. The results exhibit gout patients have higher abundance of *Prevotella*,  
16 *Fusobacterium* spp. and *Bacteroides* spp., whereas healthy controls have higher abundance of  
17 Enterobacteriaceae spp., butyrate-producing species, including *Roseburia* spp., *Butyrivibrio* spp.  
18 and *Coprococcus* spp. and anti-inflammatory *Faecalibacterium prausnitzii*. Functional analysis  
19 shows gut microbiome of gout patients have higher potential for fructose, mannose metabolism and  
20 lipid A biosynthesis, but lower potential for urate degradation and SCFAs production.  
21 Enterobacteriaceae spp. may contribute to urate degradation and provide immunostimulatory effect  
22 in healthy controls. A disease classifier based on gut microbiota shows positive performance in the  
23 discovery and validation cohorts (93.03% and 89.13% accuracy, respectively). The effect of  
24 uric-acid-lowering and anti-inflammatory drugs on the gut microbiome is mild. Integrative analyses  
25 of four additional diseases (obesity, type 2 diabetes, ankylosing spondylitis and rheumatoid arthritis)  
26 indicates gout seems to be more similar to autoimmune diseases than metabolic diseases. This work

1 demonstrates an altered gut microbiota might influence the development of gout and provides new  
2 insights into the diagnosis and treatment of the disease.

3

## 4 **1. Introduction**

5 Gout is an inflammatory arthritis disease that primarily involves the joints and is considered one of  
6 the risk factors for hypertension and cardiovascular disease<sup>[1]</sup>. Gout is more common in males than  
7 females<sup>[2]</sup>, with a rise in prevalence due to changes in diet and lifestyle<sup>[3]</sup>. In the US, the prevalence  
8 of gout was 3.9% in 2007-2008<sup>[4]</sup>. In the UK, the prevalence of gout was 2.49% in 2012, which was  
9 a 63.9% increase compared to that in 1997<sup>[5]</sup>. In eastern China, the prevalence of gout was  
10 approximately 1.14% in 2008, whereas gout was regarded as a rare disease in 1980<sup>[6]</sup>.

11 Gout is now known to be mainly caused by an abnormal increase in uric acid and the  
12 crystallization of monosodium urate (MSU) crystals. In healthy people, there are two ways to  
13 excrete uric acid from the body: the kidney, which is responsible for discharging 70% of uric acid,  
14 and the intestine, which excretes the remaining 30%<sup>[7]</sup>. In addition, diet and lifestyle, the two  
15 important risk factors of gout, have been reported to be able to influence the gut microbiota<sup>[8]</sup>.  
16 Recent studies have shown that the gut microbiota is associated with various human diseases,  
17 including obesity (OB)<sup>[9]</sup>, hypertension<sup>[10]</sup>, ankylosing spondylitis (AS)<sup>[11]</sup> and rheumatoid arthritis  
18 (RA)<sup>[12]</sup> and may play an important role in the development of those diseases. Nevertheless, the  
19 association between gut microbiota and gout is still poorly understood. A previous study had  
20 reported that gut microbiota was correlated with gout. In gout patients, the abundance of  
21 *Bacteroides caccae* and *Bacteroides xylanisolvens* was significantly enriched, whereas the  
22 abundance of *Faecalibacterium prausnitzii* and *Bifidobacterium pseudocatenulatum* was decreased.  
23 The gut microbial functions, including purine metabolism, were also altered in gout patients<sup>[13]</sup>.  
24 Another study combining microbiome and metabolome analysis showed differences in metabolites  
25 and microbial taxa between gout patients and healthy individuals. Opportunistic pathogens, such as  
26 those in *Bacteroides*, Porphyromonadaceae, *Rhodococcus*, *Erysipelatoclostridium* and

1 Anaerolineaceae were increased in gout patients<sup>[14]</sup>. These results indicated that gut microbiota was  
2 associated with gout and may participate in its development. However, both studies had a small  
3 sample size (68 and 52, respectively) without or with only a few validation samples. Additionally,  
4 the interaction between gut microbiota and gout associated clinical parameters, and the response of  
5 gut microbiota to therapeutic interventions were not described in these studies. Therefore, additional  
6 studies with larger discovery and validation cohorts, together with more detailed characterization of  
7 patient clinical parameters, are needed to provide a more comprehensive understanding the  
8 underlying impact of gut microbiota in gout.

9 In this study, we designed a two-stage metagenomics study in 188 Chinese people consisting  
10 of a discovery cohort of 77 gout patients and 63 healthy controls as well as a validation cohort of 25  
11 gout patients and 23 healthy controls. We compared the gut microbiome between groups and  
12 identified the microbial biomarkers related to gout. A disease classifier based on gut microbiota was  
13 constructed for gout. Comparison with four additional diseases revealed that gout seemed to be  
14 more similar to autoimmune diseases than metabolic diseases.

15

## 16 **2. Results**

### 17 **2.1 Intestinal microbial dysbiosis in gout patients**

18 To investigate the gut microbiota profiles in gout patients, we performed metagenomic shotgun  
19 sequencing of 307 fecal samples from 188 individuals. There were 140 samples in the discovery  
20 cohort, including 77 gout patients and 63 healthy controls, and 48 samples in the validation cohort,  
21 including 25 gout patients and 23 healthy controls. In addition, 70, 40 and 9 fecal samples were  
22 collected from gout patients in the discovery cohort longitudinally at three time points (weeks 2, 4  
23 and 24), to assess the effect of therapeutic intervention on the gut microbiome in gout. All patients'  
24 information is summarized in Table S1 (Supporting Information). Importantly, statistical analysis of  
25 clinical parameters (**Table 1**) showed that differences mainly focused on the inflammation, kidney  
26 and gout-associated indices, such as C-reactive protein (CRP), erythrocyte sedimentation rate (ESR),

1 serum creatinine (SCr), estimated glomerular filtration rate (eGFR), serum uric acid (SUA), visual  
2 analog scale (VAS), arthralgia, joint swelling scores (JSS) and joint tenderness scores (JTS).  
3 Moreover, we observed an elevated BMI in gout patients, consistent with the view that gout was  
4 associated with OB<sup>[15]</sup>. Furthermore, gout-associated indices declined after receiving medication,  
5 suggesting the relief of gout symptoms. Gout patients and healthy controls showed no significant  
6 difference on specific dietary habits, including drinking, probiotics/prebiotics and completely  
7 vegetable-based diet.

8 We carried out rarefaction analysis by random sampling 100 times with replacement to assess  
9 the metagenomic gene counts in discovery samples (healthy control, n = 63; gout patient, n = 77).  
10 As the sample size increased, both curves tended to be flat with few new genes detected, suggesting  
11 that the sequencing data were sufficient (**Figure 1A**). We found that gout patients had lower  
12 microbial gene richness and diversity than healthy controls (gene number,  $P = 0.0024$ ; Shannon  
13 index,  $P = 0.0016$ ; Wilcoxon rank-sum test; Figure 1B, C). Beta diversity analysis indicated a  
14 significantly greater dissimilarity in microbial community between samples in disease and healthy  
15 control groups than those within each group ( $P = 0.0069$ , Wilcoxon rank-sum test; Figure 1D).

16 To better understand the gut microbial differences between gout patients and healthy controls  
17 and to identify microbial taxa associated with gout, all genes were annotated, and microbial relative  
18 abundances at different taxonomic levels were obtained. In discovery cohort, a total of 18 phyla,  
19 400 genera and 861 species were identified, respectively. The 5 predominant phyla were  
20 Bacteroidetes (52.65%), Firmicutes (15.24%), Proteobacteria (3.40%) Fusobacteria (0.68%) and  
21 Actinobacteria (0.63%) (Figure S1A, Supporting Information). At the genus level, genera with an  
22 average relative abundance more 1% were *Bacteroides* (34.89%), *Prevotella* (9.68%),  
23 *Faecalibacterium* (3.07%), *Roseburia* (1.89%), *Alistipes* (1.86%), *Parabacteroides* (1.67%),  
24 *Blautia* (1.20%), *Escherichia* (1.18%) and *Megamonas* (1.07%) (Figure S1B, Supporting  
25 Information). These 9 genera accounted for more than 50% of gut microbial composition.  
26 Considering that enterotype has been reported to be related to diseases<sup>[10]</sup>, we investigated

1 enterotype of each sample in discovery cohort. We identified two distinct enterotypes in 140  
2 samples (enterotype1, 99 samples; enterotype2, 41 samples) (Figure S2A, Supporting Information)  
3 at the genus level, each containing healthy controls and gout patients. Predominant genera in  
4 enterotype1 and enterotype2 were *Bacteroides* ( $P = 1.96e-18$ , Wilcoxon rank-sum test) and  
5 *Prevotella* ( $P = 2.67e-20$ , Wilcoxon rank-sum test) (Figure S2B, Supporting Information),  
6 respectively. Moreover, no significant difference of the two enterotypes distribution between gout  
7 patients and healthy controls was observed ( $P = 0.8541$ , Fisher's exact test). And all clinical indices  
8 except SCr ( $P = 0.0118$ , Wilcoxon rank-sum test) also showed no significant difference between  
9 individuals of two enterotypes (Figure S2C, Supporting Information). These findings suggested that  
10 there was no significant association between enterotype and gout. By examining the top 5 principal  
11 components (PC), the first PC was significantly correlated with enterotype (false discovery rate  
12 (FDR) =  $1.53e-20$ , Wilcoxon rank-sum test), whereas the second was significantly correlated with  
13 gout (FDR =  $2.25e-05$ , Wilcoxon rank-sum test; Figure 1E), which indicated that gout status was  
14 one of determining factors of gut microbial composition in our study. Similar results were obtained  
15 at functional gene level (PC3, FDR =  $1.01e-06$ ; Figure S2D, E, Supporting Information). Though  
16 BMI was significantly different between two groups, none of the top 5 PCs were significantly  
17 correlated with it (Table S3, Supporting Information). Furthermore, no genus was detected to be  
18 significantly correlated with BMI using MaAsLin analysis ( $q$  value  $< 0.25$ ). These results indicated  
19 that BMI had limited influence on gut microbiota in our study.

20 Wilcoxon rank-sum test was used to detect the significantly different bacteria between two  
21 groups. At the phylum level, Bacteroidetes (FDR = 0.022, Wilcoxon rank-sum test) and  
22 Fusobacteria (FDR = 0.0183, Wilcoxon rank-sum test) were enriched in gout patients, whereas  
23 Proteobacteria (FDR = 0.0183, Wilcoxon rank-sum test) were enriched in healthy controls (Figure  
24 1G). Furthermore, we observed that the ratio of Bacteroidetes to Firmicutes (two main phyla of the  
25 human gut microbes) was higher in gout patients than in healthy controls ( $P = 0.0098$ , Wilcoxon  
26 rank-sum test; Figure 1F), which suggested that the gut microbial composition of gout patients had

1 changed. A total of 110 genera and 223 species were significantly different in abundance between  
2 the two groups (FDR < 0.05, Wilcoxon rank-sum test; Figure 1H, I and Tables S4, 5, Supporting  
3 Information). All differential *Bacteroides* spp. (3 species), *Prevotella* spp. (13 species) and  
4 *Fusobacterium* spp. (4 species) were enriched in gout patients. *Bacteroides* has been reported to be  
5 significantly higher in the gut microbiota of gout patients and may be associated with  
6 inflammation<sup>[13]</sup>. Similarly, increased *Prevotella*-mediated gut mucosal inflammation leads to  
7 systemic diseases<sup>[16]</sup>. *Prevotella copri* (FDR = 0.0169, Wilcoxon rank-sum test) was observed to be  
8 increased in untreated rheumatoid arthritis<sup>[17]</sup> and displayed the ability to increase inflammation in a  
9 mouse model of colitis<sup>[17]</sup>. *Fusobacterium* spp., especially *Fusobacterium nucleatum* (FDR =  
10 0.0035, Wilcoxon rank-sum test), have been demonstrated to correlate with colorectal cancer (CRC)  
11 and were often considered a pathogen. They can generate a proinflammatory microenvironment that  
12 is conducive to the development of CRC<sup>[18]</sup>. *Lactobacillus* spp. (15 species) and *Streptococcus* spp.  
13 (11 species) were also higher in gout patients. By contrast, butyrate-producing species ( $P < 0.05$ ,  
14 FDR < 0.1, Wilcoxon rank-sum test), such as *Roseburia* spp. (3 species), *Coprococcus* spp. (3  
15 species), *Eubacterium* spp. (3 species), *Faecalibacterium prausnitzii* and butyrate-producing  
16 bacterium SS3/4 were enriched in healthy controls. These species were reported to have  
17 anti-inflammatory effects<sup>[19,20]</sup>. Some species belonging to Enterobacteriaceae, including  
18 *Escherichia* spp. (2 species), *Klebsiella* spp. (7 species), *Enterobacter* spp. (9 species) and  
19 *Citrobacter* spp. (7 species), were reported to degrade uric acid<sup>[21]</sup>. These species were higher in  
20 healthy controls, possibly helping reduce uric acid accumulation in the healthy controls.

21

## 22 **2.2 Functional characterization of the gut microbiome in gout patients**

23 To investigate the functional differences in the microbiome between gout patients and healthy  
24 controls in discovery cohort, a KEGG Orthology (KO) profile was constructed according to the  
25 reference gene catalogue annotation. In total, 7,289 KOs were identified in all samples (n = 259), of  
26 which 2,666 KOs (2,260 and 406 KOs enriched in healthy controls and gout patients, respectively)

1 were differentially enriched between the two groups (FDR < 0.05, Wilcoxon rank-sum test; Table  
2 S6, Supporting Information). KOs profile was used to perform PCA, and the top two principal  
3 components were significantly different between gout patients and healthy controls (PC1,  $P =$   
4 0.0008; PC2,  $P = 0.0043$ , Wilcoxon rank-sum test; Figure S3A, B, Supporting Information), which  
5 indicated that the microbial function of gout patients was altered.

6 To further explore the functional changes, the different KEGG pathways and modules between  
7 the two groups were identified (Figure S3C, D and Table S7, Supporting Information). The gut  
8 microbial function of gout patients displayed higher potential for carbohydrate metabolism,  
9 including fructose and mannose metabolism, starch and sucrose metabolism, amino sugar and  
10 nucleotide sugar metabolism, the citrate cycle and 3 modules of glycolysis/gluconeogenesis. Low  
11 carbohydrates was showed to reduce SUA and gout attacks<sup>[22]</sup>. The increase in fructose was  
12 reported to be associated with the risk of gout<sup>[23]</sup>, and a previous study showed that a rapid infusion  
13 of fructose could raise the level of SUA<sup>[24]</sup>. Concurrently, we observed increased potential for  
14 lipopolysaccharide (LPS) biosynthesis and lipid A biosynthesis (Figure S4A, B, Supporting  
15 Information). LPS was reported to be a common component of gram-negative bacteria and could be  
16 an endotoxin associated with inflammation and disease<sup>[25]</sup>. In the two groups, *Bacteroides* and  
17 *Prevotella* were the major contributors to lipid A biosynthesis. We note that lipid A biosynthesis  
18 genes from Proteobacteria (FDR = 0.017, Wilcoxon rank-sum test) species were enriched in healthy  
19 controls (Figure S4C, Supporting Information), including *Escherichia coli* (FDR = 0.005, Wilcoxon  
20 rank-sum test) and *Klebsiella pneumonia* (FDR = 0.014, Wilcoxon rank-sum test). In contrast, the  
21 samples of healthy controls showed higher potential for cell motility, containing bacterial  
22 chemotaxis and flagellar assembly, similar to the previous results observed in AS<sup>[11]</sup> and type 2  
23 diabetes (T2D)<sup>[26]</sup>, and such mechanisms may be beneficial for health. Previous studies indicated  
24 that the gut microbiota may take part in purine metabolism<sup>[13,14]</sup>. Consistently, we observed that  
25 healthy controls had a higher potential for urate degradation (FDR < 0.05, Wilcoxon rank-sum test;  
26 **Figure 2A, B**) and the major contributor was Enterobacteriaceae spp. (FDR < 0.05, Spearman's

1 rank correlation; Figure 2C and Table S8, Supporting Information) in which *Klebsiella* showed the  
2 highest Spearman's rank correlation (mean Spearman's rank correlation of 5 urate degradation KOs  
3 was 0.53). Meanwhile, SUA was negatively correlated with Enterobacteriaceae spp., especially  
4 *Klebsiella* ( $r = -0.23$ ,  $P = 0.0057$ , Spearman's rank correlation; Figure 2D and Figure S5, Supporting  
5 Information), which indicated that the gut microbiome had the ability to degrade uric acid. To verify  
6 this result, we investigated the distribution of KOs involved in urate degradation among 1,135  
7 species (Figure S6 and Table S9, Supporting Information). Proteobacteria (706 species) was the  
8 predominant contributor, and the largest family was Enterobacteriaceae (119 species), which is  
9 consistent with our results. As SCFAs have been reported to be linked to healthy gut microbiota, we  
10 analyzed the abundance of genes encoding key enzymes involving SCFAs production. We found  
11 that the genes responsible for propionate and butyrate were significantly different between gout  
12 patients and healthy controls, whereas genes for acetate did not show any difference (Figure S7,  
13 Supporting Information).

14

### 15 **2.3 Metagenomic species (MGS) associated with gout**

16 To further explore the features of the gout-associated microbiome and the correlations between  
17 clinical indices and the microbial alteration in gout patients, 178,233 differentially enriched genes  
18 (60,511 and 117,722 enriched in healthy controls ( $n = 63$ ) and gout patients ( $n = 77$ ), respectively)  
19 were clustered into 3,831 co-abundance gene groups (CAGs) based on the co-variation of relative  
20 abundance across all samples using the Canopy-based algorithm. CAGs with fewer than 50 genes  
21 were filtered, and the remaining 134 were termed metagenomic species (MGSs) for follow-up  
22 analysis (Table S10, Supporting Information). All MGSs were significantly different between the  
23 two groups, where 73 MGSs were more abundant in gout patients, and the remaining 61 MGSs  
24 were more abundant in healthy controls ( $FDR < 0.05$ , Wilcoxon rank-sum test). Notably, only 24 of  
25 the healthy control-enriched MGSs could be annotated as specific species compared with 53 of gout  
26 patients. Subsequently, we carried out a co-occurrence network analysis for all MGSs based on

1 SparCC correlation coefficients (pseudo  $P < 0.05$ ; Figure S8, Supporting Information). We observed  
2 more intensive and sophisticated interactions between gout-enriched and gout-depleted MGSs than  
3 within each group.

4 Permutational analysis of variance (PERMANOVA) was performed to assess the influence of  
5 clinical indices on MGSs and we observed that the disease status was the strongest factor (lowest  
6  $P$ -values in PERMANOVA; Table S11, Supporting Information). Other clinical indices such as  
7 SUA, CRP, eGFR, ESR and SCr also showed correlation with microbial alteration. Canonical  
8 coordinate analysis (CCA) showed that gout-enriched and control-enriched MGSs were clearly  
9 separated (Figure S9, Supporting Information). All clinical indices were associated with  
10 gout-enriched MGSs except for eGFR, and gout status was the main factor affecting the composition  
11 of the gut microbiome, followed by SUA. Spearman's rank correlation coefficient analysis was  
12 conducted for physiological indices and MGSs (FDR  $< 0.05$ ; Figure S10 and Table S12, Supporting  
13 Information). A high level of SUA is a major feature of gout patients. We found SUA was positively  
14 correlated with many gout-enriched MGSs but negatively correlated with many control-enriched  
15 MGSs. Interestingly, we observed that almost all *Bacteroides* spp. enriched in gout patients were  
16 correlated with SUA, whereas no association was found between *Bacteroides* spp. enriched in  
17 healthy controls and SUA. The indicator of inflammation CRP showed a similar result. SCr  
18 displayed positive correlation with several gout-enriched MGSs. An increase in the level of SCr  
19 suggests impaired renal function. BMI, age, eGFR, ESR and urea displayed a correlation with a small  
20 number of MGSs. Furthermore, we carried out PCoA and PERMANOVA on the MGS level for gout  
21 related phenotypes, including gout duration, VAS, JSS, JST and tophi of gout patient samples (Figure  
22 S11A-E and Table S11, Supporting Information), only tophi significantly separated the gout patient  
23 samples. In a subsequent step, significantly different MGSs ( $P < 0.05$ , Wilcoxon rank-sum test) were  
24 detected between gout patients with and without tophi. Control-enriched MGSs, such as *F. prausnitzii*,  
25 were higher in patients without tophi, whereas gout-enriched MGSs, especially *Bacteroides* spp.,  
26 were enriched in patients with tophi (Figure S11F, Supporting Information).

1

## 2 **2.4 Gut microbiota-based gout classification**

3 The distinct microbial differences between gout patients and healthy controls prompted us to  
4 investigate if the gut microbiome had the potential to distinguish gout patients from healthy controls.  
5 A disease classifier was constructed based on the MGS profile using a random forest model in the  
6 discovery cohort (healthy control, n = 63; gout patient, n = 77), and the corresponding receiver  
7 operating characteristic (ROC) curve was constructed. All MGSs were entered into the disease  
8 classifier, and 16 MGSs were selected as biomarkers to establish a classification model with a  
9 fivefold cross-validation approach (**Figure 3A-C** and Table S13, Supporting Information). Notably,  
10 the area under the receiver operating curve (AUC) reached 0.93. We found that only three MGS  
11 biomarkers were enriched in healthy controls, whereas others were enriched in gout patients. Five  
12 of the top ten most important MGSs for classification, including Gout-107, *Bacteroides sp. (100)*,  
13 *Megamonas sp. (102)*, *Blautia sp. (92)*, *Parabacteroides distasonis (74)*, had the lowest *P* value. To  
14 evaluate the accuracy of this classification model, an independent cohort consisting of 25 gout  
15 patients and 23 healthy controls was used for validation, and the AUC value was 0.89 (Figure 3D, E  
16 and Table S14, Supporting Information). In addition, 16 MGS biomarkers were used to distinguish  
17 patients from healthy controls in the AS (n = 211), RA (n = 169), T2D (n = 268) and OB (n = 200)  
18 cohorts (Figure 3F). The AUC was lower than in the gout cohort, suggesting that these 16 MGS  
19 biomarkers were gout-specific. These results indicated that gout-associated microbial biomarkers  
20 have the potential to be used as a novel and noninvasive diagnostic strategy for gout.

21

## 22 **2.5 Effect of uric-acid-lowering and anti-inflammatory drugs intervention on the gut** 23 **microbiome**

24 Previous studies of metagenomics have demonstrated the potential influence of therapeutic drugs on  
25 the microbiota<sup>[27]</sup>. Here, we assessed the effect of drug intervention on intestinal microbiota in this  
26 gout cohort. We conducted a 2-week (2W, n = 61), 4-week (4W, n = 38), and 24-week (24W, n = 7)

1 treatment intervention for gout patients. Most of the treated gout patients received  
2 uric-acid-lowering and anti-inflammatory drugs intervention (Figure S12A and Table S1,  
3 Supporting Information). We excluded 9, 2, and 2 fecal samples from the 2W, 4W and 24W groups,  
4 respectively, from this analysis, because these patients did not use any drugs or used other  
5 non-conventional drugs (the sample size was too small), such as NaHCO<sub>3</sub> and traditional Chinese  
6 medicine. Subsequently, we compared their microbial composition with untreated gout patients (n =  
7 77) and healthy controls (n = 63). The gene numbers, Shannon index and PCoA showed that the gut  
8 microbial composition was significantly different between healthy controls and gout patients. ( $P <$   
9 0.05, Wilcoxon rank-sum test; Figure S12B, C, Supporting Information). However, no significant  
10 difference was found between before and after treatment in gout patients, and PERMANOVA  
11 analysis also supported this result (Table S15, Supporting Information). Considering that the  
12 unmatched sample size at different time points may influence the results, we selected five gout  
13 patients whose fecal samples were collected at all the time points for further analysis. Interestingly,  
14 we observed samples of 24W treatment and found that the gut microbial composition was further  
15 away from other time points and more similar to samples of healthy controls (Figure S12D,  
16 Supporting Information). This may indicate that the effect of uric-acid-lowering and  
17 anti-inflammatory drugs on the gut microbiome was modest and partially restored the gut  
18 microbiome over time, similar to disease-modifying anti-rheumatic drugs (DMARD)<sup>[12]</sup>.  
19 Allopurinol, colchicine, celecoxib and etoricoxib had no impact on human gut bacteria *in vitro* in a  
20 previous report<sup>[28]</sup>. Benzbromarone has been shown to inhibit gut bacterial growth *in vitro*<sup>[28]</sup> Our  
21 results suggested a more complicated relationship between these drugs and gut microbiota *in vivo*.  
22 However, due to the insufficient data in our study, the mechanisms of specific drugs on the gut  
23 microbiome needs to be further studied in intervention trials of larger cohorts.

24

## 25 **2.6 Comparison of the gut microbiome between gout and other diseases**

26 To better understand the potential role of intestinal microbes in gout, we compared the gut microbial

1 composition of gout with AS, RA, T2D and OB, of which the first two are autoimmune diseases,  
2 and the last two are metabolic diseases. However, as the comorbidity data is not available, we  
3 cannot exclude that there might be individuals with a combination of multiple diseases. From the  $P$   
4 value distribution, we inferred that the degree of gut microbial dysbiosis was intermediate in gout  
5 patients when compared with the other four diseases, with OB being more severe and RA being  
6 least severe (**Figure 4A**). The number of significant microbiota markers also supported this  
7 conclusion (Figure 4B). Very few significant microbiota markers (Wilcoxon rank-sum test, FDR <  
8 0.05, only 6 genes FDR less than 0.05, Figure 4B) were detected in RA patients. Surprisingly, the  
9 gut microbial composition of gout patients was closer to AS patients (Figure 4C and Figure S13A,  
10 Supporting Information), and they shared 40 different species. A significant enrichment of  
11 *Acidaminococcus fermentans*, *Streptococcus anginosus*, *Megasphaera elsdenii* and *Megasphaera*  
12 *micronuciformis* was observed in all diseases except RA. In addition, *Streptococcus* spp. also had a  
13 higher abundance in patients. For gut microbial functions, gout patients were also more similar to  
14 autoimmune diseases RA and AS (Figure 4D and Figure S13B, Supporting Information). The  
15 functions of oxidative phosphorylation; alanine, aspartate and glutamate metabolism; carbon  
16 fixation pathways in prokaryotes; and carbapenem biosynthesis were enriched in gout, AS and RA  
17 patients, whereas the functions of bacterial chemotaxis and flagellar assembly were depleted in  
18 these three diseases and T2D. Except for fructose and mannose metabolism and *S. aureus* infection,  
19 no differential function was shared among gout, OB or T2D. Collectively, these results indicated  
20 that the gut microbiome of gout patients was more similar to autoimmune diseases than metabolic  
21 diseases.

22

### 23 **3. Discussion**

24 In this metagenome-wide association study, we observed that the gut microbiome of gout patients  
25 was substantially different from healthy controls. Specifically, in gout patients, the abundance of  
26 *Bacteroides* spp. and *Prevotella* was enriched, while the abundance of Enterobacteriaceae spp. and

1 SCFA-producing bacteria was depleted. Analysis of microbial functions demonstrated that gout  
2 patients had an enhanced ability to metabolize fructose and mannose and synthesize LPS and lipid  
3 A but a reduced ability to degrade purine and generate SCFAs. These alterations in gut microbial  
4 species and functions may contribute to the development of gout (**Figure 5**). Furthermore, based on  
5 the identified differential biomarkers in the discovery cohort, we constructed a disease classifier to  
6 explore the potential of gut microbiota for the diagnosis of gout, and it performed well in the  
7 validation cohort. Finally, we observed that the gut microbiome of gout was more similar to AS and  
8 RA when compared with T2D and OB, which suggested that gout was possibly an autoimmune  
9 disease, at least from the perspective of the gut microbiome.

10 The ability to degrade uric acid is widely distributed among the Enterobacteriaceae<sup>[21]</sup>. Our  
11 analysis results also supported this notion (Figure 2). Enterobacteriaceae species had the ability to  
12 utilize purines or grow with uric acid as a source of nitrogen and carbon<sup>[29]</sup>. We observed that the  
13 relative abundance of *K. pneumoniae*, *E. coli*, *Enterobacter cloacae* and *Citrobacter freundii*, which  
14 belong to Enterobacteriaceae, decreased in gout patients and negatively correlated with SUA level,  
15 suggesting that Enterobacteriaceae may take part in uric acid degradation in the gut and maintain  
16 the stability of uric acid in human body. Additionally, increasing the intake of sugar-sweetened soft  
17 drinks and fructose was associated with a higher risk of gout<sup>[30,31]</sup>. Fructose may increase *de novo*  
18 purine biosynthesis to contribute to the production of uric acid<sup>[32]</sup>. In this study, we observed some  
19 microbial functions associated with fructose metabolism, including starch and sucrose metabolism,  
20 and fructose and mannose metabolism were enriched in gout patients, which may help increase the  
21 production of uric acid, thus increasing the risk of gout.

22 Previous studies have shown that bacterial components including LPS play a crucial role in  
23 immune homeostasis<sup>[33]</sup>. Moreover, it is reported that LPS is involved in gouty inflammation<sup>[34]</sup>.  
24 The Lipid A is the key component of LPS for host immune recognition. The lipid A with six acyl  
25 chains produced by Enterobacteriaceae is able to induce endotoxin tolerance, which is a healthy  
26 immunostimulatory effect. However, this effect can be inhibited by the lipid A with 4 or 5 acyl

1 chains produced by *Bacteroides* or *Prevotella*<sup>[35,36]</sup>. In our results, lipid A biosynthesis genes from  
2 *Bacteroides* and *Prevotella* enriched in gout patients may reduce the healthy immune stimulating  
3 effect of Enterobacteriaceae, which may contribute to the inflammation in gout. Furthermore, the  
4 Enterobacteriaceae species *E. coli* has been reported to suppress arthritis, possibly through LPS in  
5 animal models<sup>[37,38]</sup>. Interestingly, *Klebsiella*, belonging to Enterobacteriaceae, has also been shown  
6 to be depleted in autoimmune diseases, including RA<sup>[12]</sup> and AS<sup>[11]</sup>. These results indicated that  
7 Enterobacteriaceae may have a regulatory effect on arthritic diseases.

8 Dietary fiber can be fermented by the gut microbiota to form SCFAs that have  
9 anti-inflammatory effects. Accumulated evidence have indicated that SCFAs are associated with  
10 inflammatory diseases, including inflammatory bowel disease (IBD)<sup>[39]</sup>, asthma<sup>[40]</sup> and allergies<sup>[41]</sup>.  
11 SCFAs also affect gout inflammation. Butyrate was reported to be able to suppress acute gout  
12 arthritis by inhibiting histone deacetylases (HDACs) and decrease MSU-induced production of  
13 IL-1 $\beta$ , IL-6 and IL-8<sup>[42]</sup>. Furthermore, a previous reported that butyrate biosynthesis in the gut was  
14 depleted in gout patients<sup>[13]</sup>. Acetate was also shown to promote resolution of the inflammatory  
15 response induced by MSU in a model of gout mice<sup>[43]</sup>. Butyrate-producing bacteria<sup>[44-46]</sup>, such as  
16 *Roseburia* spp., *Butyrivibrio* spp., *Coprococcus* spp., *F. prausnitzii*, and butyrate-producing  
17 bacterium (Table S5, Supporting Information), and butyrate-producing enzymes (Figure S7,  
18 Supporting Information) were enriched in healthy controls in this study, suggesting that butyrate  
19 may play an important role in remitting gout inflammation. Cell mobility, including flagellar  
20 assembly and bacterial chemotaxis, was enriched in healthy controls, and flagellin could induce  
21 epithelial cells to secrete antimicrobial peptides that help control the growth of opportunistic  
22 pathogens. All these microbial functions may help maintain a healthy gut environment in healthy  
23 controls, thereby reducing the risk of gout.

24 Based on the alteration of gut microbial composition, we constructed a classifier to  
25 discriminate healthy controls and gout patients and obtained high accuracy (0.93 and 0.89 for the  
26 discovery and validation cohorts, respectively). It confirmed that microbial biomarkers we found in

1 our study had a high specificity which suggested the possible application of these biomarkers in  
2 disease diagnosis and treatment. Previous reports also found that microbial biomarkers could be  
3 used as a novel and noninvasive method for diagnosing gout<sup>[13]</sup> or other diseases<sup>[47,48]</sup>. However, the  
4 microbiota-based disease prediction model was restricted by geography<sup>[49]</sup>. Future studies should  
5 collect a large geographic scale to construct a common model for predicting gout status. Moreover,  
6 gut microbiome alteration may occur earlier than clinical symptoms<sup>[10]</sup>. A cohort with high SUA but  
7 no gout could be added to investigate the dynamic of the gut microbiome.

8 For gout treatment, although SUA decreased after taking a uric-acid-lowering drug, a previous  
9 report recommended that drug treatment should be continued for 6 months for gout patients without  
10 tophi<sup>[50]</sup>. Consistently, we found that the composition of the gut microbiota was not affected by  
11 drugs treatment over a short time period but partially restored over 6 months (Figure S12,  
12 Supporting Information). This finding implied that gut microbiota may play an additional key role  
13 in treatment, and the modest regulation of gut microbiota might be related to gout relapse after  
14 discontinuing the current mainstream drugs. Considering the sample size of gout treatment was  
15 small, a larger cohort study is expected to be performed to validate our results in the future. In view  
16 of the influence of intestinal microbes on gout, gout may be prevented, and the risk can be reduced  
17 through the restoration of the gut by improving lifestyle and diet or early intervention with drugs.

18 We observed that the gut microbiota of gout was more similar to immune diseases in terms of  
19 the microbial composition and functions (Figure 4), although RA was not clustered together with  
20 gout and AS at the species level. Although we are unsure how these species and functions affect the  
21 host immune system, the similarities in gut microbiota suggested that they may play a common role  
22 in the development of immunity. However, because of the small sample size and limitations of  
23 statistical analysis methods, a larger sample size and further analysis are needed for evaluation and  
24 validation.

25

## 26 **4. Conclusion**

1 Our results demonstrated the significant differences of gut microbial composition and functions  
2 between gout patients and healthy controls and found gut microbiota may take part in the  
3 development of gout. The high performance of our disease classification model suggested the  
4 potential of gut microbiota for the diagnosis of gout in clinical practice. However, sophisticated  
5 models considering confounding factors such as geography should be constructed and validated in  
6 larger independent cohorts. Further studies should use longitudinal samples and animal models to  
7 investigate the dynamics of gout and determine the causality of gut microbiota and gout. Our study  
8 also suggested that gout patients may improve their condition by adjusting the gut microbiota  
9 through changing lifestyle or regulating diet.

10

## 11 **5. Experimental Section**

12 *Subjects:* This study was approved by the Medical Ethics Committee of the Second Affiliated  
13 Hospital of Guangzhou University of Chinese Medicine (B2016-103-01). All of the participants  
14 provided written informed consent. Patients were diagnosed with gout as determined by the 2015  
15 ACR/EULAR classification criteria<sup>[51]</sup> and suitability for the treatment in this clinical trial. Patients  
16 and healthy controls had to meet the following criteria: 1) 15-69 years of age; 2) no antibiotics and  
17 glucocorticoid use within 3 months and 1 month, respectively; 3) no gastrointestinal diseases, such  
18 as gastrointestinal surgery, Crohn's disease, ulcerative colitis, or acute diarrhea; 4) no history of  
19 severe, progressive or uncontrolled cardiac, hepatic, renal, mental, or hematological disease; and 5)  
20 no history of drug abuse.

21 Between May 2016 and September 2018, we recruited 102 male acute gout patients (SUA,  
22  $543.1 \pm 128.4 \mu\text{mol/L}$ , data are mean  $\pm$  s.d.) and 86 age-matched male healthy controls (SUA,  $358.6$   
23  $\pm 52.3 \mu\text{mol/L}$ , data are mean  $\pm$  s.d.) from the Chinese population for this study (Table S1,  
24 Supporting Information). After collecting fecal samples at baseline, gout patients were treated with  
25 uric-acid-lowering (benzbromarone, allopurinol, febuxostat) and anti-inflammatory drugs  
26 (colchicine, celecoxib, etoricoxib, betamethasone, voltaren), and fecal samples were collected after

1 treatment for 2 weeks (n = 70), 4 weeks (n = 40), and 24 weeks (n = 9). A total of 307 fecal samples  
2 were collected from the Second Affiliated Hospital of Guangzhou University of Chinese Medicine  
3 and frozen at -80 °C until transport. Food frequency questionnaire was collected from all  
4 participants to assess dietary differences between two groups.

5 Blood samples were collected and frozen at -80 °C until analysis. Serum uric acid (SUA),  
6 serum creatinine (SCr), urea nitrogen, C-reactive protein (CRP), and estimated glomerular filtration  
7 rate (eGFR) were measured using the Cobas 8000 modular analyzer (Roche, Switzerland), and  
8 erythrocyte sedimentation rate (ESR) was measured using the Test-1 analyzer (Alifax, Italy).

9 *DNA extraction and library construction:* The collected fecal specimens were centrifuged at  
10 12000 × g at room temperature for 5 min and the supernatant was discarded. 200 mg pellet was  
11 weighted from each sample and used for total bacterial DNA extraction with the E.Z.N.A Stool  
12 DNA Kit (OMEGA Bio-tek, USA) according the manufacturer's instructions. The quality of DNA  
13 was analyzed using Qubit (Invitrogen, USA) and 1% agarose gel electrophoresis. The detail of  
14 DNA library construction was described in Supporting Information file. The final DNA library was  
15 determined the average insert size using the Agilent 2100 Bioanalyzer (Agilent Technologies, USA)  
16 and quantified by ABI StepOnePlus Real-Time PCR system (Applied Biosystems, USA).

17 *Metagenomic sequencing:* Paired-end metagenomic sequencing was performed on the Illumina  
18 HiSeq 4000 platform with an insert size of 350 bp and paired-end (PE) reads of 150 bp for each  
19 sample. After removing adaptors and low quality and ambiguous bases from the raw reads, the  
20 remaining reads were aligned to human genome reference (hg19) by SOAPaligner (v2.22,  
21 parameters: -m 280 -x 420 -r 1 -l 32 -s 75 -c 0.9) to remove human host DNA contamination. The  
22 average rate of host contamination was  $0.52 \pm 2.06\%$ . Finally, 2,768.76 Gb of high-quality PE reads  
23 for the 307 samples were acquired with an average of 9.02 Gb per sample (Table S2, Supporting  
24 Information). The metagenomic shotgun sequencing data for all samples have been deposited in the  
25 CNGB Nucleotide Sequence Archive (CNSA) under accession code CNP0000284.

26 *Construction of gene, phylum, genus, species and KO profiles:* The clean reads were aligned to

1 the 11,446,577 genes in the reference gene catalogue<sup>[52]</sup> using SOAPaligner (v2.22, parameters: -m  
2 200 -x 1000 -r 2 -v 13 -l 32 -s 75 -c 0.95), and  $72.79 \pm 2.89\%$  reads ( $n = 307$ , data are mean  $\pm$  s.d.)  
3 were mapped. The gene abundance profile was calculated as previously described<sup>[11]</sup>. After  
4 removing genes detected in less than 10% of the discovery samples ( $n = 259$ ), 1,564,977 genes  
5 remained. To improve the gene taxonomy annotation result, genes were aligned to the National  
6 Center for Biotechnology Information (NCBI) microbial reference genomes (including 5,847  
7 microbial genomes, v20171114) and NT database (v20170924) using BLAT (v.36) and Megablast  
8 (v2.2.26) with default parameters, respectively. The cutoff was at least 70% overlap of query and 65%  
9 identity for phylum, 85% identity for genus, and 95% identity for species<sup>[53]</sup>. The relative  
10 abundance of phyla, genera, species and KOs were calculated from the relative abundance of their  
11 respective genes using previously published methods<sup>[53]</sup>.

12 *Rarefaction curve, gene counts and biodiversity analysis:* Rarefaction analysis was performed  
13 to assess the gene richness in the healthy controls and gout patients. For a given number of samples,  
14 we performed random sampling 100 times in the cohort with replacement and estimated the total  
15 number of genes that could be identified from these samples by the Chao2 richness estimator<sup>[54]</sup>.  
16 The total gene counts in each sample were calculated as previously described<sup>[55]</sup>. The alpha  
17 diversity and beta diversity were estimated by the Shannon index and Bray-Curtis distance,  
18 respectively. To adjust for the effect of the various data sizes among the samples, these analyses  
19 were based on the gene profile that was randomly sampled to 11 million (M) matched reads.

20 *Identification of enterotypes:* The enterotypes of each sample in discovery cohort were  
21 identified by the PAM clustering algorithm using the relative abundance of genus.  
22 Calinski-Harabaze (CH) index was used to assess the optimal number of clusters<sup>[56]</sup>. Finally, PCA  
23 was performed using Jensen-Shannon distance to cluster the samples.

24 *Gene functional analysis:* Differentially enriched KEGG pathways and modules were  
25 identified according to reporter score<sup>[57]</sup> from the Z-scores of individual KEGG Orthologies (KOs)  
26 (KEGG database release 79). Pathways or modules were considered significantly different if the

1  $|\text{reporter score}| > 1.65$ , corresponding to 95% confidence to a normal distribution.

2 For analysis of urate degradation genes, species that contain at least one urate degradation gene  
3 in the KEGG database (v79) were collected, and the corresponding 16S sequences were  
4 downloaded from the Ribosomal Database Project (RDP, set14). Then, 16S sequences were aligned  
5 by PyNAST (v1.2.2), and trees were constructed by fasttree (v2.1.7). The tree and heat map were  
6 visualized in iTOL<sup>[58]</sup>.

7 The protein sequences of SCFA-producing enzymes were obtained from the NCBI database<sup>[59]</sup>.  
8 Genes in the reference gut microbiome gene catalogue were aligned to these sequences using  
9 BLASTP (v2.2.26, best hit with e-value  $< 1e-5$ , identity  $> 70\%$  and coverage  $> 70\%$ ), and their  
10 relative abundance were calculated as KOs.

11 *Co-abundance gene groups (CAGs)*: A total of 178,233 genes differed in relative abundance  
12 between healthy controls and gout patients in the discovery cohort (Wilcoxon rank-sum test,  $P <$   
13  $0.01$ , FDR  $< 0.05$ ) from 1,564,977 genes (present at least 10% of the samples in discovery cohort).  
14 These marker genes were clustered into CAGs based on their relative abundance variation across all  
15 samples using the Canopy-based algorithm with default parameters<sup>[60]</sup>. In brief, a gene was  
16 randomly selected as a ‘seed’ and this and other genes with similar abundance profiles were  
17 clustered as a CAG if Pearson correlation coefficient of those genes was more than 0.9. In order to  
18 avoid the profile being too sparse, relative abundance of gene in the third quartiles of each group  
19 was selected as a cluster relative abundance. Finally, 3,831 CAGs (contain at least two genes) were  
20 obtained. Clusters including at least 50 genes remained for further analysis and were called  
21 metagenomic species (MGSs). MGSs were assigned to a species, genus or phylum if more than 50%  
22 of their genes had the same level of assignment<sup>[60]</sup>. The relationship between the abundance of each  
23 MGS was according to correlation calculated by SparCC<sup>[61]</sup>, and the co-occurrence network was  
24 visualized with Cytoscape (v3.4.0).

25 *Canonical correspondence analysis (CCA)*: CCA was performed on the MGS abundance  
26 profile to assess the effect of each of the indices listed. The plot was generated by R (v3.4.0, vegan

1 package).

2 *PERMANOVA on the influence of phenotype*: Permutation multivariate analysis of variance  
3 (PERMANOVA) was performed on the MGS abundance profile to assess the impact from each of  
4 the clinical indices listed<sup>[62]</sup>. We used Bray-Curtis distance and 9,999 permutations in R (3.4.0,  
5 vegan package).

6 *MGS-based classifier*: We constructed a classifier to discriminate healthy controls and gout  
7 patients based on a random forest model (randomForest 4.6-14 package) using the profile of MGS  
8 relative abundance<sup>[62]</sup>. A fivefold cross-validation approach was used to evaluate the predictive  
9 model. We calculated the average cross-validation error curves and the minimum error in the  
10 averaged curve, plus the standard deviation at that point was used as the cutoff to filter the  
11 predictive model. The ones containing the smallest number of MGSs among all sets with an error  
12 below the cutoff were chosen as a diagnosis classifier. Then, the probability of gout was calculated  
13 based on this optimal set, and the receiver operating characteristic (ROC) for both the discovery  
14 cohort and validation cohort were drawn using the pROC package<sup>[63]</sup>.

15 *Comparison of the gut microbiota between gout and other diseases*: Published metagenomic  
16 sequencing data of ankylosing spondylitis (AS) (SRP100575 and ERP005860; 114 healthy controls  
17 and 97 patients with AS)<sup>[11]</sup>, rheumatoid arthritis (RA) (ERP006678; 74 healthy controls and 95  
18 patients with RA)<sup>[12]</sup>, type 2 diabetes (T2D) (SRP008047 and SRP011011; 185 healthy controls and  
19 183 patients with T2D)<sup>[26]</sup> and obesity (OB) (ERP013562; 105 healthy controls and 95 patients with  
20 OB)<sup>[9]</sup> were downloaded from NCBI or the European Bioinformatics Institute (EBI). All sequencing  
21 reads were mapped to the same reference gut gene catalogue and analyzed for gene function by the  
22 same pipeline as the gout cohort. To avoid methodological differences, taxonomic analyses of genes  
23 for all cohorts was based on the reference gene catalogue taxa result, including the gout cohort.

24 *Statistical analysis*: All statistical analyses were performed by R (v3.4.0). Differential relative  
25 abundance of genes, taxa and KOs were detected by Wilcoxon rank-sum with an adjusted *P* value  
26 (corrected by the Benjamini-Hochberg)  $< 0.05$ . Enrichment in healthy controls or gout patients was

1 determined according to the higher mean rank-sum. The correlations between the relative  
2 abundance of differentially genera and KOs involved to urate degradation KOs, the relative  
3 abundance of MGSs and clinical indices were calculated by Spearman's rank correlation coefficient  
4 and visualized by heatmap in R using the 'pheatmap' package. The association between species and  
5 BMI was calculated by MaAsLin with default parameters<sup>[64]</sup>.

## 7 **Availability of data and material**

8 The metagenomic shotgun sequencing data for all samples have been deposited in the CNGB  
9 Nucleotide Sequence Archive (CNSA) under accession code CNP0000284. Other data that support  
10 the findings of this study are available within the paper and its Supplementary Information files or  
11 from the corresponding author upon reasonable request.

## 13 **Conflict of Interest**

14 The authors declare no conflict of interest.

## 16 **Acknowledgements**

17 This study was supported by grants from 1010 project from Guangdong Provincial Hospital of  
18 Chinese Medicine (No. YN10101906), the grants from Guangdong Provincial Hospital of Chinese  
19 Medicine (No. YN2018ML08, No. YN2018ZD06), Guangzhou Municipal Science and Technology  
20 supporting project (No. 201710010076), Guangdong Science and Technology Department (No.  
21 2016A020226041) as well as Guangzhou science and technology program key projects  
22 (NO.201604020005).

## 24 **Authors' contributions**

25 R. Huang, Q. Huang and S. Sun conceived and designed this project; R. Huang, S. Sun and Q. Gao  
26 managed the project; Y. Chu, J. Li, M. Wang, X. He, X. Chen and X. Wu collected the samples and

1 performed clinical diagnosis; Y. Huang, X. Xie and P. Wang contributed to metagenomic data  
2 analysis; Y. Huang and L. Liang finished the KEGG database urate degradation genes analyses. Y.  
3 Huang, X. Xie and P. Wang wrote the manuscript; R. Huang, S. Sun and Q. Gao provided guidance  
4 for manuscript preparation. All authors contributed to the revision of the paper.

5

## 6 References

- 7 [1] A. So, B. Thorens, *J. Clin. Invest.* **2010**, *120*, 1791.
- 8 [2] P. Richette, T. Bardin, *Lancet* **2010**, *375*, 318.
- 9 [3] K. G. Saag, H. Choi, *Arthritis Res. Ther.* **2006**, *8 Suppl 1*, S2.
- 10 [4] Y. Zhu, B. J. Pandya, H. K. Choi, *Arthritis Rheum.* **2011**, *63*, 3136.
- 11 [5] C.-F. Kuo, M. J. Grainge, C. Mallen, W. Zhang, M. Doherty, *Ann. Rheum. Dis.* **2015**, *74*,  
12 661.
- 13 [6] Z. Miao, C. Li, Y. Chen, S. Zhao, Y. Wang, Z. Wang, X. Chen, F. Xu, F. Wang, R. Sun, J.  
14 Hu, W. Song, S. Yan, C.-Y. Wang, *J. Rheumatol.* **2008**, *35*, 1859.
- 15 [7] L. B. Sorensen, D. J. Levinson, *Nephron* **1975**, *14*, 7.
- 16 [8] J. Wang, H. Jia, *Nat. Rev. Microbiol.* **2016**, *14*, 508.
- 17 [9] R. Liu, J. Hong, X. Xu, Q. Feng, D. Zhang, Y. Gu, J. Shi, S. Zhao, W. Liu, X. Wang, H. Xia,  
18 Z. Liu, B. Cui, P. Liang, L. Xi, J. Jin, X. Ying, X. Wang, X. Zhao, W. Li, H. Jia, Z. Lan, F.  
19 Li, R. Wang, Y. Sun, M. Yang, Y. Shen, Z. Jie, J. Li, X. Chen, H. Zhong, H. Xie, Y. Zhang,  
20 W. Gu, X. Deng, B. Shen, X. Xu, H. Yang, G. Xu, Y. Bi, S. Lai, J. Wang, L. Qi, L. Madsen,  
21 J. Wang, G. Ning, K. Kristiansen, W. Wang, *Nat. Med.* **2017**, *23*, 859.
- 22 [10] J. Li, F. Zhao, Y. Wang, J. Chen, J. Tao, G. Tian, S. Wu, W. Liu, Q. Cui, B. Geng, W. Zhang,  
23 R. Weldon, K. Auguste, L. Yang, X. Liu, L. Chen, X. Yang, B. Zhu, J. Cai, *Microbiome*  
24 **2017**, *5*, 14.
- 25 [11] C. Wen, Z. Zheng, T. Shao, L. Liu, Z. Xie, E. Le Chatelier, Z. He, W. Zhong, Y. Fan, L.  
26 Zhang, H. Li, C. Wu, C. Hu, Q. Xu, J. Zhou, S. Cai, D. Wang, Y. Huang, M. Breban, N. Qin,

- 1 S. D. Ehrlich, *Genome Biol.* **2017**, *18*, 142.
- 2 [12] X. Zhang, D. Zhang, H. Jia, Q. Feng, D. Wang, D. Liang, X. Wu, J. Li, L. Tang, Y. Li, Z.  
3 Lan, B. Chen, Y. Li, H. Zhong, H. Xie, Z. Jie, W. Chen, S. Tang, X. Xu, X. Wang, X. Cai, S.  
4 Liu, Y. Xia, J. Li, X. Qiao, J. Y. Al-Aama, H. Chen, L. Wang, Q. Wu, F. Zhang, W. Zheng,  
5 Y. Li, M. Zhang, G. Luo, W. Xue, L. Xiao, J. Li, W. Chen, X. Xu, Y. Yin, H. Yang, J. Wang,  
6 K. Kristiansen, L. Liu, T. Li, Q. Huang, Y. Li, J. Wang, *Nat. Med.* **2015**, *21*, 895.
- 7 [13] Z. Guo, J. Zhang, Z. Wang, K. Y. Ang, S. Huang, Q. Hou, X. Su, J. Qiao, Y. Zheng, L.  
8 Wang, E. Koh, H. Danliang, J. Xu, Y. K. Lee, H. Zhang, *Sci. Rep.* **2016**, *6*, 20602.
- 9 [14] T. Shao, L. Shao, H. Li, Z. Xie, Z. He, C. Wen, *Front. Microbiol.* **2017**, *8*, 268.
- 10 [15] R. Roubenoff, M. J. Klag, L. A. Mead, K. Y. Liang, A. J. Seidler, M. C. Hochberg, *J. Am.*  
11 *Med. Assoc.* **1991**, *266*, 3004.
- 12 [16] E. Abdullah, A. Idris, A. Saparon, *ARPN J. Eng. Appl. Sci.* **2017**, *12*, 3218.
- 13 [17] J. U. Scher, A. Sczesnak, R. S. Longman, N. Segata, C. Ubeda, C. Bielski, T. Rostron, V.  
14 Cerundolo, E. G. Pamer, S. B. Abramson, C. Huttenhower, D. R. Littman, *Elife* **2013**, *2*,  
15 e01202.
- 16 [18] A. D. Kostic, E. Chun, L. Robertson, J. N. Glickman, C. A. Gallini, M. Michaud, T. E.  
17 Clancy, D. C. Chung, P. Lochhead, G. L. Hold, E. M. El-Omar, D. Brenner, C. S. Fuchs, M.  
18 Meyerson, W. S. Garrett, *Cell Host Microbe* **2013**, *14*, 207.
- 19 [19] P. B. Eckburg, E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill,  
20 K. E. Nelson, D. A. Relman, *Science (80-. )*. **2005**, *308*, 1635.
- 21 [20] Z. Tamanai-Shacoori, I. Smida, L. Bousarghin, O. Loreal, V. Meuric, S. B. Fong, M.  
22 Bonnaure-Mallet, A. Jolivet-Gougeon, *Future Microbiol.* **2017**, *12*, 157.
- 23 [21] G. D. Vogels, C. Van der Drift, *Bacteriol. Rev.* **1976**, *40*, 403.
- 24 [22] P. H. Dessen, E. A. Shipton, A. E. Stanwix, B. I. Joffe, J. Ramokgadi, *Ann. Rheum. Dis.*  
25 **2000**, *59*, 539.
- 26 [23] J. Jamnik, S. Rehman, S. Blanco Mejia, R. J. de Souza, T. A. Khan, L. A. Leiter, T. M. S.

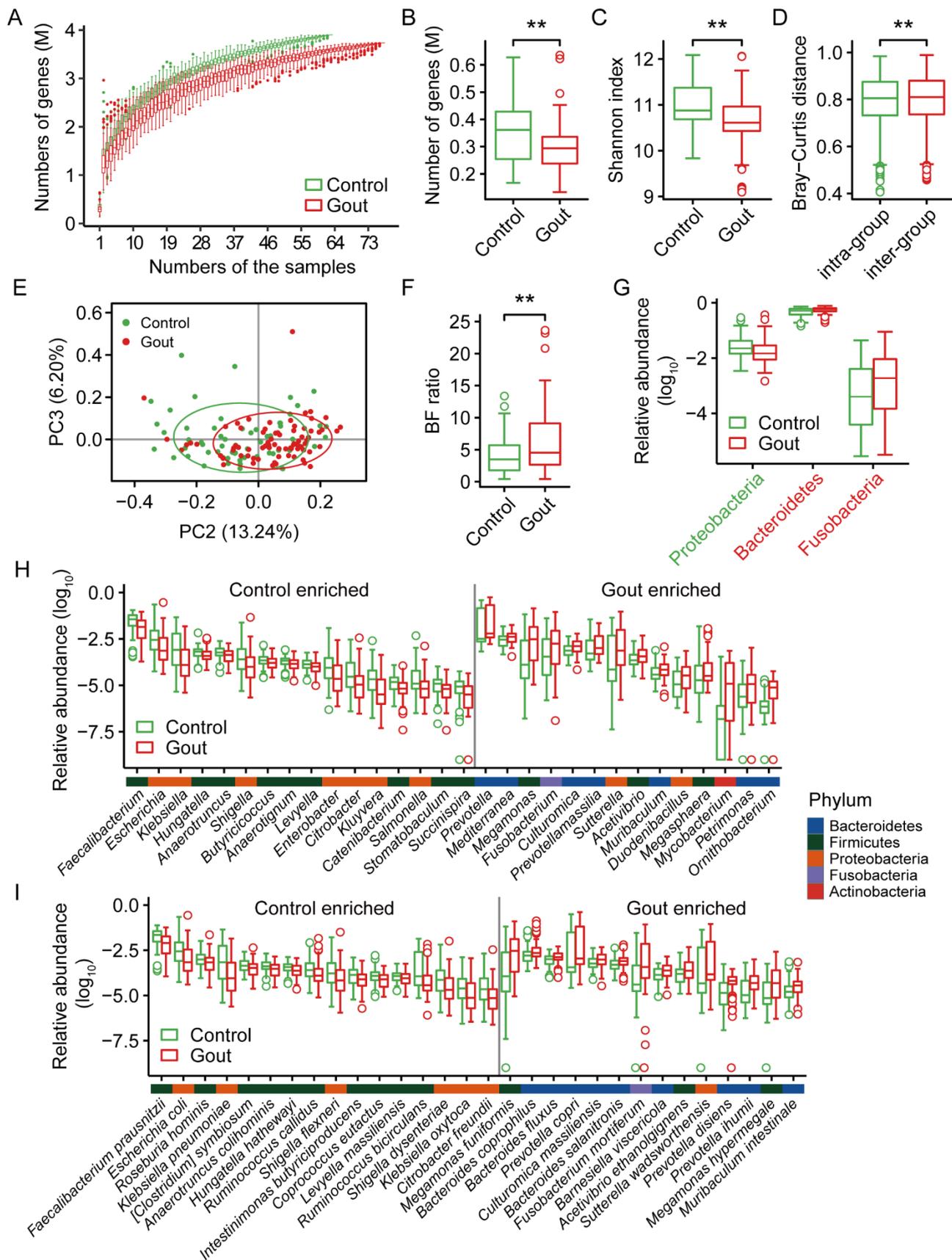
- 1 Wolever, C. W. C. Kendall, D. J. A. Jenkins, J. L. Sievenpiper, *BMJ Open* **2016**, *6*, e013191.
- 2 [24] R. G. Narins, J. S. Weisberg, A. R. Myers, *Metabolism* **1974**, *23*, 455.
- 3 [25] J. R. Marchesi, D. H. Adams, F. Fava, G. D. A. Hermes, G. M. Hirschfield, G. Hold, M. N.
- 4 Quraishi, J. Kinross, H. Smidt, K. M. Tuohy, L. V. Thomas, E. G. Zoetendal, A. Hart, *Gut*
- 5 **2016**, *65*, 330.
- 6 [26] J. Qin, Y. Li, Z. Cai, S. Li, J. Zhu, F. Zhang, S. Liang, W. Zhang, Y. Guan, D. Shen, Y. Peng,
- 7 D. Zhang, Z. Jie, W. Wu, Y. Qin, W. Xue, J. Li, L. Han, D. Lu, P. Wu, Y. Dai, X. Sun, Z. Li,
- 8 A. Tang, S. Zhong, X. Li, W. Chen, R. Xu, M. Wang, Q. Feng, M. Gong, J. Yu, Y. Zhang, M.
- 9 Zhang, T. Hansen, G. Sanchez, J. Raes, G. Falony, S. Okuda, M. Almeida, E. LeChatelier, P.
- 10 Renault, N. Pons, J.-M. Batto, Z. Zhang, H. Chen, R. Yang, W. Zheng, S. Li, H. Yang, J.
- 11 Wang, S. D. Ehrlich, R. Nielsen, O. Pedersen, K. Kristiansen, J. Wang, *Nature* **2012**, *490*,
- 12 55.
- 13 [27] Z. Jie, H. Xia, S.-L. Zhong, Q. Feng, S. Li, S. Liang, H. Zhong, Z. Liu, Y. Gao, H. Zhao, D.
- 14 Zhang, Z. Su, Z. Fang, Z. Lan, J. Li, L. Xiao, J. Li, R. Li, X. Li, F. Li, H. Ren, Y. Huang, Y.
- 15 Peng, G. Li, B. Wen, B. Dong, J.-Y. Chen, Q.-S. Geng, Z.-W. Zhang, H. Yang, J. Wang, J.
- 16 Wang, X. Zhang, L. Madsen, S. Brix, G. Ning, X. Xu, X. Liu, Y. Hou, H. Jia, K. He, K.
- 17 Kristiansen, *Nat. Commun.* **2017**, *8*, 845.
- 18 [28] L. Maier, M. Pruteanu, M. Kuhn, G. Zeller, A. Telzerow, E. E. Anderson, A. R. Brochado, K.
- 19 C. Fernandez, H. Dose, H. Mori, K. R. Patil, P. Bork, A. Typas, *Nature* **2018**, *555*, 623.
- 20 [29] M. A. Rouf, R. F. Lomprey, Jr., *J. Bacteriol.* **1968**, *96*, 617.
- 21 [30] H. K. Choi, G. Curhan, *BMJ* **2008**, *336*, 309.
- 22 [31] H. K. Choi, W. Willett, G. Curhan, *JAMA* **2010**, *304*, 2270.
- 23 [32] B. T. Emmerson, *Ann. Rheum. Dis.* **1974**, *33*, 276.
- 24 [33] S. M. Man, *Nat. Rev. Gastroenterol. Hepatol.* **2018**, *15*, 721.
- 25 [34] S. R. Kingsbury, P. G. Conaghan, M. F. McDermott, *J. Inflamm. Res.* **2011**, *4*, 39.
- 26 [35] T. Vatanen, A. D. Kostic, E. D’Hennezel, H. Siljander, E. A. Franzosa, M. Yassour, R.

- 1 Kolde, H. Vlamakis, T. D. Arthur, A. M. Hämäläinen, A. Peet, V. Tillmann, R. Uibo, S.  
2 Mokurov, N. Dorshakova, J. Ilonen, S. M. Virtanen, S. J. Szabo, J. A. Porter, H. Lähdesmäki,  
3 C. Huttenhower, D. Gevers, T. W. Cullen, M. Knip, R. J. Xavier, *Cell* **2016**, *165*, 842.
- 4 [36] E. d’Hennezel, S. Abubucker, L. O. Murphy, T. W. Cullen, *mSystems* **2017**, *2*, e00046.
- 5 [37] O. Kohashi, Y. Kohashi, T. Takahashi, A. Ozawa, N. Shigematsu, *Arthritis Rheum.* **1986**, *29*,  
6 547.
- 7 [38] O. Kohashi, Y. Kohashi, T. Takahashi, A. Ozawa, N. Shigematsu, *Microbiol. Immunol.* **1985**,  
8 *29*, 487.
- 9 [39] K. Machiels, M. Joossens, J. Sabino, V. De Preter, I. Arijs, V. Eeckhaut, V. Ballet, K. Claes,  
10 F. Van Immerseel, K. Verbeke, M. Ferrante, J. Verhaegen, P. Rutgeerts, S. Vermeire, *Gut*  
11 **2014**, *63*, 1275.
- 12 [40] A. N. Thorburn, C. I. McKenzie, S. Shen, D. Stanley, L. Macia, L. J. Mason, L. K. Roberts,  
13 C. H. Y. Wong, R. Shim, R. Robert, N. Chevalier, J. K. Tan, E. Mariño, R. J. Moore, L.  
14 Wong, M. J. McConville, D. L. Tull, L. G. Wood, V. E. Murphy, J. Mattes, P. G. Gibson, C.  
15 R. Mackay, *Nat. Commun.* **2015**, *6*, 7320.
- 16 [41] A. Trompette, E. S. Gollwitzer, K. Yadava, A. K. Sichelstiel, N. Sprenger, C. Ngom-Bru, C.  
17 Blanchard, T. Junt, L. P. Nicod, N. L. Harris, B. J. Marsland, *Nat. Med.* **2014**, *20*, 159.
- 18 [42] M. C. P. Cleophas, T. O. Crişan, H. Lemmers, H. Toenhake-Dijkstra, G. Fossati, T. L. Jansen,  
19 C. A. Dinarello, M. G. Netea, L. A. B. Joosten, *Ann. Rheum. Dis.* **2016**, *75*, 593.
- 20 [43] A. T. Vieira, I. Galvão, L. M. Macia, É. M. Sernaglia, M. A. R. Vinolo, C. C. Garcia, L. P.  
21 Tavares, F. A. Amaral, L. P. Sousa, F. S. Martins, C. R. Mackay, M. M. Teixeira, *J. Leukoc.*  
22 *Biol.* **2017**, *101*, 275.
- 23 [44] A. Rivière, M. Selak, D. Lantin, F. Leroy, L. De Vuyst, *Front. Microbiol.* **2016**, *7*, 979.
- 24 [45] P. Van Den Abbeele, C. Belzer, M. Goossens, M. Kleerebezem, W. M. De Vos, O. Thas, R.  
25 De Weirdt, F. M. Kerckhof, T. Van De Wiele, *ISME J.* **2013**, *7*, 949.
- 26 [46] P. Louis, H. J. Flint, *FEMS Microbiol. Lett.* **2009**, *294*, 1.

- 1 [47] J. Yu, Q. Feng, S. H. Wong, D. Zhang, Q. Yi Liang, Y. Qin, L. Tang, H. Zhao, J. Stenvang,  
2 Y. Li, X. Wang, X. Xu, N. Chen, W. K. K. Wu, J. Al-Aama, H. J. Nielsen, P. Kiilerich, B. A.  
3 H. Jensen, T. O. Yau, Z. Lan, H. Jia, J. Li, L. Xiao, T. Y. T. Lam, S. C. Ng, A. S. L. Cheng,  
4 V. W. S. Wong, F. K. L. Chan, X. Xu, H. Yang, L. Madsen, C. Datz, H. Tilg, J. Wang, N.  
5 Brünner, K. Kristiansen, M. Arumugam, J. J. Y. Sung, J. Wang, *Gut* **2017**, *66*, 70.
- 6 [48] R. Loomba, V. Seguritan, W. Li, T. Long, N. Klitgord, A. Bhatt, P. S. Dulai, C. Caussy, R.  
7 Bettencourt, S. K. Highlander, M. B. Jones, C. B. Sirlin, B. Schnabl, L. Brinkac, N. Schork,  
8 C. H. Chen, D. A. Brenner, W. Biggs, S. Yooseph, J. C. Venter, K. E. Nelson, *Cell Metab.*  
9 **2017**, *25*, 1054.
- 10 [49] Y. He, W. Wu, H. M. Zheng, P. Li, D. McDonald, H. F. Sheng, M. X. Chen, Z. H. Chen, G.  
11 Y. Ji, Z. D. X. Zheng, P. Mujagond, X. J. Chen, Z. H. Rong, P. Chen, L. Y. Lyu, X. Wang, C.  
12 Bin Wu, N. Yu, Y. J. Xu, J. Yin, J. Raes, R. Knight, W. J. Ma, H. W. Zhou, *Nat. Med.* **2018**,  
13 *24*, 1532.
- 14 [50] T. Neogi, *N. Engl. J. Med.* **2011**, *364*, 443.
- 15 [51] T. Neogi, T. L. T. A. Jansen, N. Dalbeth, J. Fransen, H. R. Schumacher, D. Berendsen, M.  
16 Brown, H. Choi, N. L. Edwards, H. J. E. M. Janssens, F. Lioté, R. P. Naden, G. Nuki, A.  
17 Ogdie, F. Perez-Ruiz, K. Saag, J. A. Singh, J. S. Sundry, A. K. Tausche, J. Vaquez-Mellado,  
18 S. A. Yarows, W. J. Taylor, *Ann. Rheum. Dis.* **2015**, *74*, 1789.
- 19 [52] J. K. Goodrich, E. R. Davenport, M. Beaumont, M. A. Jackson, R. Knight, C. Ober, T. D.  
20 Spector, J. T. Bell, A. G. Clark, R. E. Ley, *Cell Host Microbe* **2016**, *19*, 731.
- 21 [53] J. Li, J. Wang, H. Jia, X. Cai, H. Zhong, Q. Feng, S. Sunagawa, M. Arumugam, J. R.  
22 Kultima, E. Prifti, T. Nielsen, A. S. Juncker, C. Manichanh, B. Chen, W. Zhang, F. Levenez,  
23 J. Wang, X. Xu, L. Xiao, S. Liang, D. Zhang, Z. Zhang, W. Chen, H. Zhao, J. Y. Al-Aama, S.  
24 Edris, H. Yang, J. Wang, T. Hansen, H. B. Nielsen, S. Brunak, K. Kristiansen, F. Guarner, O.  
25 Pedersen, J. Doré, S. D. Ehrlich, P. Bork, *Nat. Biotechnol.* **2014**, *32*, 834.
- 26 [54] A. Chao, *Biometrics* **1987**, *43*, 783.

- 1 [55] E. Le Chatelier, T. Nielsen, J. Qin, E. Prifti, F. Hildebrand, G. Falony, M. Almeida, M.  
2 Arumugam, J. M. Batto, S. Kennedy, P. Leonard, J. Li, K. Burgdorf, N. Grarup, T. Jørgensen,  
3 I. Brandslund, H. B. Nielsen, A. S. Juncker, M. Bertalan, F. Levenez, N. Pons, S. Rasmussen,  
4 S. Sunagawa, J. Tap, S. Tims, E. G. Zoetendal, S. Brunak, K. Clément, J. Doré, M.  
5 Kleerebezem, K. Kristiansen, P. Renault, T. Sicheritz-Ponten, W. M. De Vos, J. D. Zucker, J.  
6 Raes, T. Hansen, P. Bork, J. Wang, S. D. Ehrlich, O. Pedersen, E. Guedon, C. Delorme, S.  
7 Layec, G. Khaci, M. Van De Guchte, G. Vandemeulebrouck, A. Jamet, R. Dervyn, N.  
8 Sanchez, E. Maguin, F. Haimet, Y. Winogradski, A. Cultrone, M. Leclerc, C. Juste, H.  
9 Blottière, E. Pelletier, D. Lepaslier, F. Artiguenave, T. Bruls, J. Weissenbach, K. Turner, J.  
10 Parkhill, M. Antolin, C. Manichanh, F. Casellas, N. Boruel, E. Varela, A. Torrejon, F.  
11 Guarner, G. Denariáz, M. Derrien, J. E. T. Van Hylckama Vlieg, P. Veiga, R. Oozeer, J.  
12 Knol, M. Rescigno, C. Brechot, C. M'Rini, A. Mérieux, T. Yamada, *Nature* **2013**, *500*, 541.
- 13 [56] M. Arumugam, J. Raes, E. Pelletier, D. Le Paslier, T. Yamada, D. R. Mende, G. R.  
14 Fernandes, J. Tap, T. Bruls, J. Batto, M. Bertalan, N. Borruel, M. Consortium, J.  
15 Weissenbach, S. D. Ehrlich, P. Bork, *Nature* **2011**, *473*, 174.
- 16 [57] K. R. Patil, J. Nielsen, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 2685.
- 17 [58] I. Letunic, P. Bork, *Nucleic Acids Res.* **2016**, *44*, W242.
- 18 [59] M. J. Claesson, I. B. Jeffery, S. Conde, S. E. Power, E. M. O'Connor, S. Cusack, H. M. B.  
19 Harris, M. Coakley, B. Lakshminarayanan, O. O'Sullivan, G. F. Fitzgerald, J. Deane, M.  
20 O'Connor, N. Harnedy, K. O'Connor, D. O'Mahony, D. van Sinderen, M. Wallace, L.  
21 Brennan, C. Stanton, J. R. Marchesi, A. P. Fitzgerald, F. Shanahan, C. Hill, R. P. Ross, P. W.  
22 O'Toole, *Nature* **2012**, *488*, 178.
- 23 [60] H. B. Nielsen, M. Almeida, A. S. Juncker, S. Rasmussen, J. Li, S. Sunagawa, D. R. Plichta, L.  
24 Gautier, A. G. Pedersen, E. Le Chatelier, E. Pelletier, I. Bonde, T. Nielsen, C. Manichanh, M.  
25 Arumugam, J.-M. Batto, M. B. Quintanilha Dos Santos, N. Blom, N. Borruel, K. S. Burgdorf,  
26 F. Boumezbear, F. Casellas, J. Doré, P. Dworzynski, F. Guarner, T. Hansen, F. Hildebrand,

- 1 R. S. Kaas, S. Kennedy, K. Kristiansen, J. R. Kultima, P. Léonard, F. Levenez, O. Lund, B.  
2 Moumen, D. Le Paslier, N. Pons, O. Pedersen, E. Prifti, J. Qin, J. Raes, S. Sørensen, J. Tap, S.  
3 Tims, D. W. Ussery, T. Yamada, MetaHIT Consortium, P. Renault, T. Sicheritz-Ponten, P.  
4 Bork, J. Wang, S. Brunak, S. D. Ehrlich, MetaHIT Consortium, *Nat. Biotechnol.* **2014**, *32*,  
5 822.
- 6 [61] J. Friedman, E. J. Alm, *PLoS Comput. Biol.* **2012**, *8*, 1.
- 7 [62] Q. Feng, S. Liang, H. Jia, A. Stadlmayr, L. Tang, Z. Lan, D. Zhang, H. Xia, X. Xu, Z. Jie, L.  
8 Su, X. Li, X. Li, J. Li, L. Xiao, U. Huber-Schönauer, D. Niederseer, X. Xu, J. Y. Al-Aama, H.  
9 Yang, J. Wang, K. Kristiansen, M. Arumugam, H. Tilg, C. Datz, J. Wang, *Nat. Commun.*  
10 **2015**, *6*, 6528.
- 11 [63] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-C. Sanchez, M. Müller, *BMC*  
12 *Bioinformatics* **2011**, *12*, 77.
- 13 [64] X. C. Morgan, T. L. Tickle, H. Sokol, D. Gevers, K. L. Devaney, D. V Ward, J. A. Reyes, S.  
14 A. Shah, N. Leleiko, S. B. Snapper, A. Bousvaros, J. Korzenik, B. E. Sands, R. J. Xavier, C.  
15 Huttenhower, *Genome Biol.* **2012**, R79.



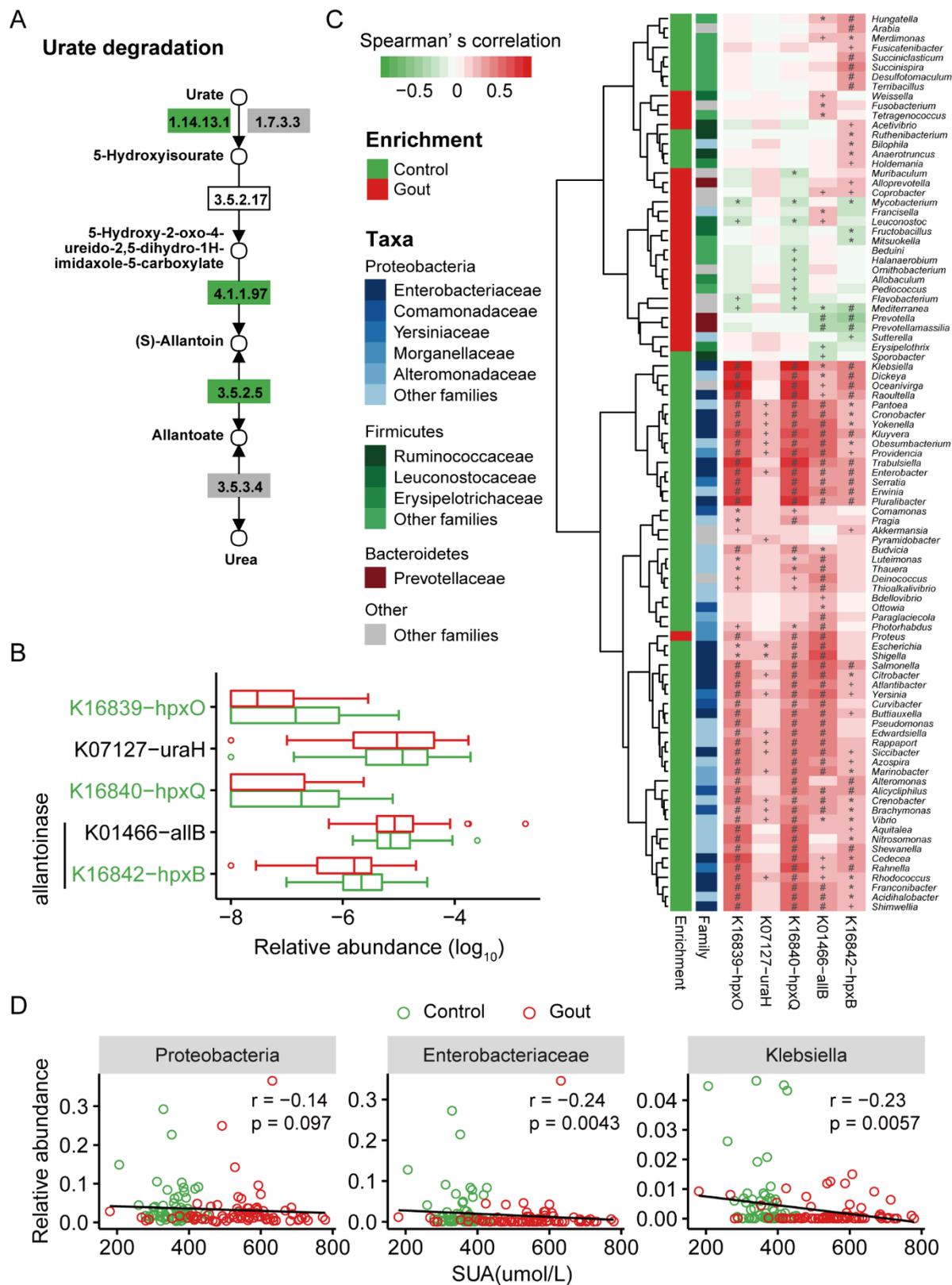
1

2 **Figure 1.** Gut microbial alterations in gout patients. (A) The gene rarefaction curves based on the

3 Chao2 values in healthy controls (n = 63) and gout patients (n = 77). (B) Box and whisker plots of

1 gene count in the healthy controls and gout patients. (C, D) Box and whisker plots of alpha diversity  
2 (Shannon index) and beta diversity (Bray Curtis distance) at the gene level. Wilcoxon rank-sum test  
3 was used to determine significance. ‘\*\*\*’  $P < 0.01$ . To exclude the influence of the various data sizes  
4 among the samples, Figures A, B, C and D were based on 11 M matched reads per individual. (E)  
5 Principle component analysis (PCA) based on the genus abundance profile. Considering PC1  
6 significantly related with enterotype, PC2 and PC3 were display. (F) The Bacteroidetes/Firmicutes  
7 ratio (Wilcoxon rank-sum test; ‘\*\*\*’,  $P < 0.01$ ). (G, H, I) The relative abundance of differential  
8 phyla (G, top 3), genera (H, top 30) and species (I, top 30) between gout patient and healthy control  
9 groups (FDR  $< 0.05$ , Wilcoxon rank-sum test). The color bar above genera or species names were  
10 colored according to the phylum.

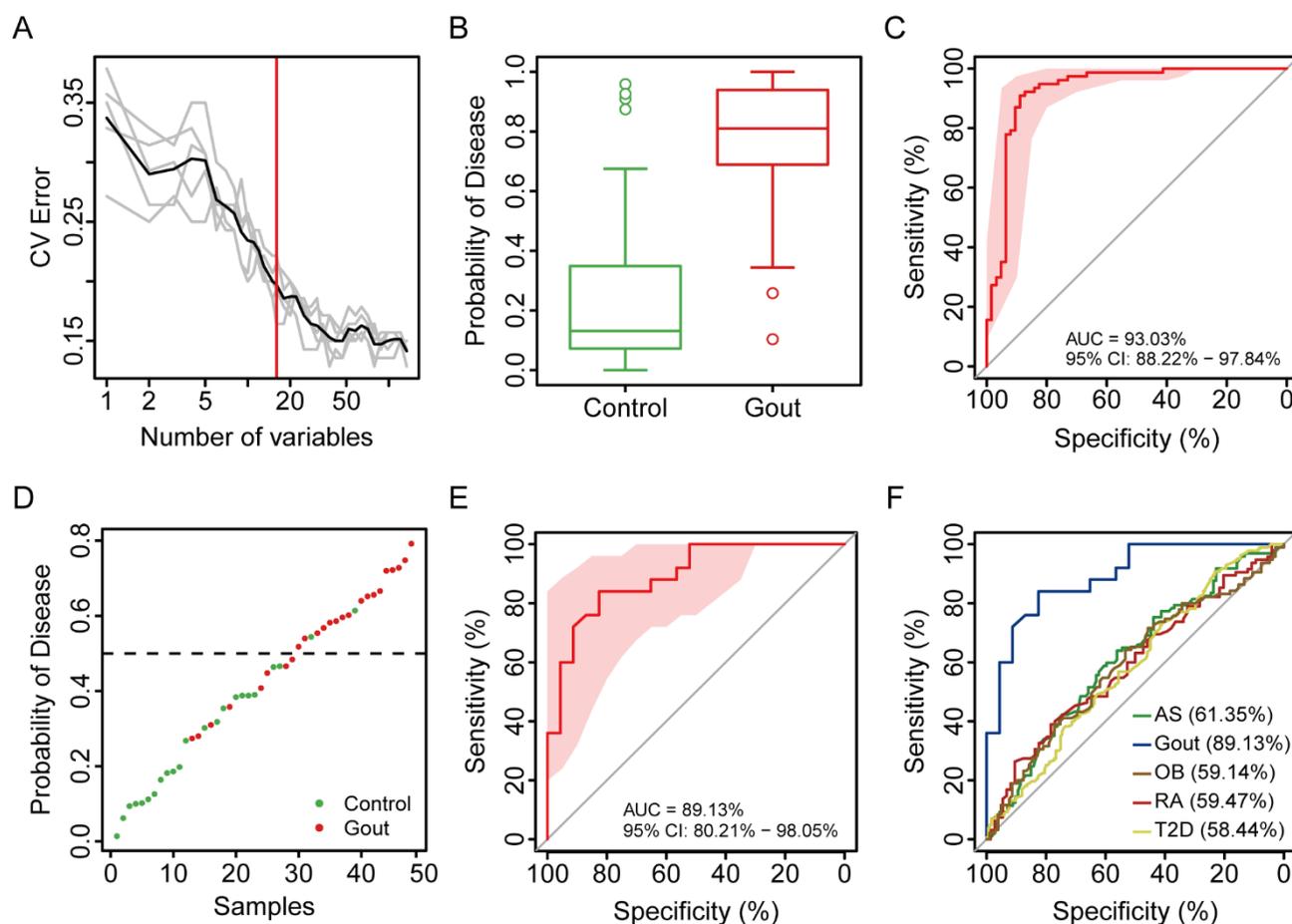
11



1

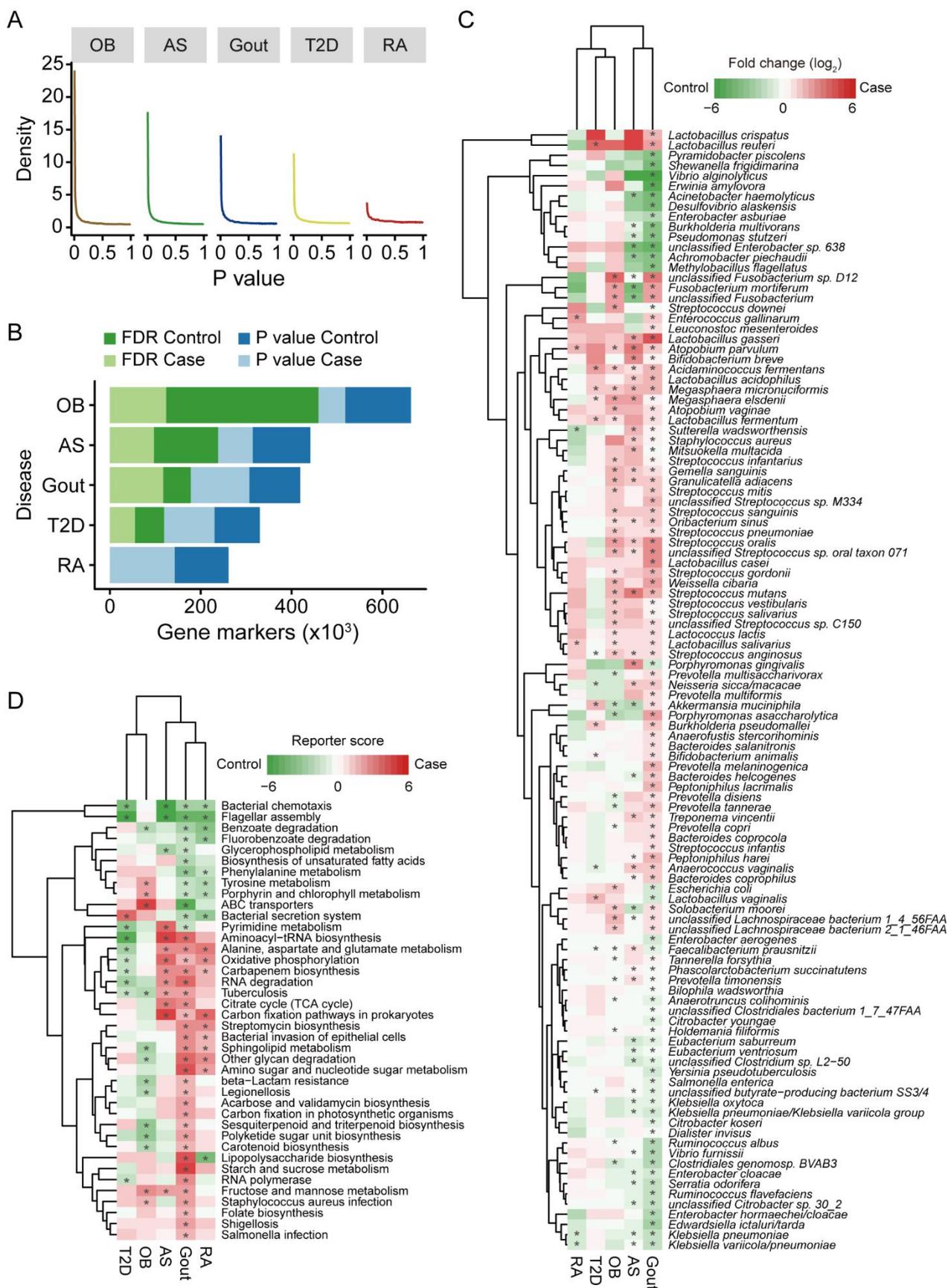
2 **Figure 2.** Gout-associated microbial gene functions related to urate degradation. (A) Module for  
 3 urate degradation. (B) Relative abundance of KOs involved in urate degradation. Significantly  
 4 enriched KOs were identified by a Wilcoxon rank-sum test, and the boxes or KO names were

1 colored according to the direction of enrichment. Green, enriched in healthy controls (FDR < 0.05).  
 2 Boxes with no color or KO names with black, no difference; boxes with gray, not detected in  
 3 samples. (C) Correlations between gout-associated genera and urate degradation-associated KOs  
 4 (red and green for positive and negative correlation, respectively). ‘+’ denotes FDR < 0.05; ‘\*’  
 5 denotes FDR < 0.01; ‘#’ denotes FDR < 0.001. (D) The associations between SUA and healthy  
 6 control-enriched species. Spearman’s rank correlation was calculated by taking the species relative  
 7 abundance and SUA content. An inverse correlation was observed between SUA and Proteobacteria,  
 8 Enterobacteriaceae and *Klebsiella*. Abbreviation: hpxO, FAD-dependent urate hydroxylase; uraH,  
 9 5-hydroxyisourate hydrolase; hpxQ, 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline  
 10 decarboxylase; allB and hpxB, allantoinase.  
 11



12  
 13 **Figure 3. Classification of gout status by relative abundance of MGS.** (A) Distribution of 5 trials  
 14 of 10-fold cross-validation error in random forest classification of gout as the number of MGS

1 increases. The model was trained using the relative abundance of the MGS (> 50 genes) in the  
2 controls and gout samples. The black curve indicates the average of the five trials (gray lines). The  
3 red line marks the number of MGSs in the optimal set. (B) The probability of gout in the  
4 cross-validation discovery samples according to the 16 MGS biomarkers. (C) Receiver operating  
5 curve (ROC) for the discovery samples. (D) Classification of the validation samples (healthy  
6 control, n = 63; gout patient, n = 77). (E) ROC for the validation samples (healthy control, n = 23;  
7 gout patient, n = 25). (F) ROCs were obtained for five cohorts using 16 gout-associated MGS  
8 biomarkers. The AUC for each disease is displayed in brackets.  
9

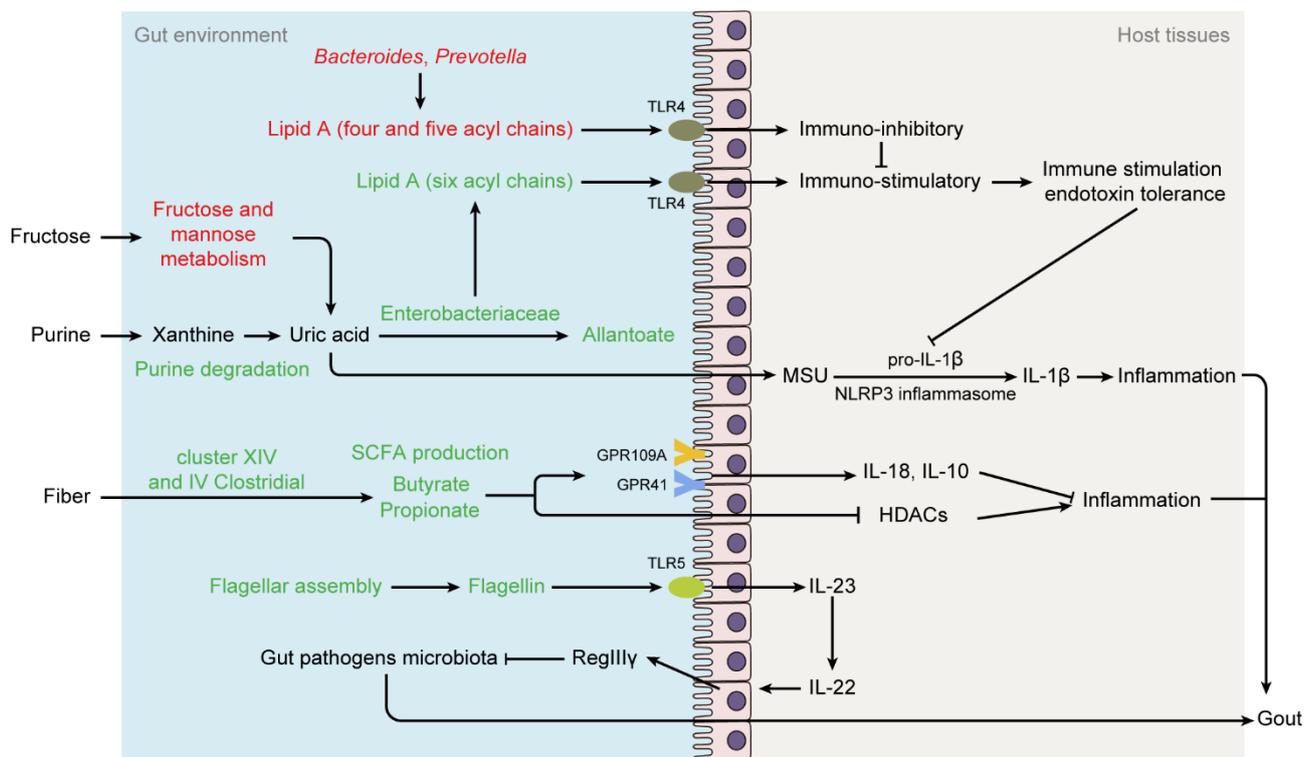


1

2 **Figure 4.** Comparison of gut microbiome and functions in different diseases. (A) The distribution  
3 of  $P$ -values in AS ( $n = 211$ ), gout ( $n = 140$ ), OB ( $n = 200$ ), RA ( $n = 169$ ) and T2D ( $n = 268$ ). (B)

1 The number of gene markers for each disease. Statistical differences between control and case of  
 2 each disease was evaluated using Wilcoxon rank-sum test and adjusted by the Benjamin-Hochberg  
 3 method. (C) Comparison of differential species in AS, gout, OB, RA and T2D. Green, enriched in  
 4 healthy controls; red, enriched in patients; ‘\*’ denotes FDR < 0.05. (D) Comparison of microbial  
 5 gene functions in AS, gout, OB, RA and T2D. Green, enriched in healthy controls; red, enriched in  
 6 patients. ‘\*’ denotes reporter score of maps > 1.65 or < -1.65.

7



8

9 **Figure 5.** Possible mechanisms of gut microbes affecting the development of gout. Green text  
 10 denotes depleted species or functions in gout patients; red text denotes enriched species or functions  
 11 in gout patients. Abbreviation: MSU, monosodium urate; TLR4, Toll-like receptor 4; TLR5,  
 12 Toll-like receptor 5; GPR41, G protein-coupled receptor 41; GPR109, G protein-coupled receptor  
 13 109A; HDAC, histone deacetylases; Th17, T-helper 17; IL-1β, interleukin-1β; IL-23, interleukin-23;  
 14 IL-22, interleukin-22.

15

16 **Table 1.** The baseline characteristics of healthy controls (n = 63) and gout patients (n = 77) in

1 discovery cohort.

Characteristics	Control; n =	Gout; n =	2W; n =	4W; n =	24W; n =	Adjusted <i>P</i> value <sup>a)</sup>	Adjusted <i>P</i> value
	63	77	70	40	9		
Age, mean ± SD	40.0 ± 12.1	39.9 ± 12.9	-	-	-	9.75E-01	9.75E-01
BMI, mean ± SD, kg/m <sup>2</sup>	23.5 ± 3.0	25.0 ± 2.5	-	-	-	6.97E-05	1.01E-04
CRP, mean ± SD, mmol/L	2.0 ± 3.2	24.5 ± 29.9	5.1 ± 7.3	4.5 ± 6.5	2.3 ± 1.9	1.93E-20	5.02E-20
eGFR, mean ± SD, ml/min	99.7 ± 15.4	90.6 ± 22.8	88.6 ± 21.0	87.1 ± 20.0	90.8 ± 14.8	1.37E-02	1.78E-02
ESR, mean ± SD, mm/h	8.8 ± 7.5	30.4 ± 24.6	18.6 ± 16.2	19.6 ± 15.7	10.1 ± 6.1	1.47E-08	2.73E-08
SCr, mean ± SD, µmol/L	82.3 ± 12.2	95.1 ± 21.2	96.5 ± 20.0	97.1 ± 22.3	93.0 ± 13.2	7.38E-06	1.20E-05
SUA, mean ± SD, µmol/L	361.4 ± 50.4	539.2 ± 118.7	400.9 ± 141.1	393.2 ± 141.3	386.8 ± 107.2	3.25E-15	7.04E-15
Urea, mean ± SD, mmol/L	5.0 ± 1.1	4.8 ± 1.5	4.6 ± 1.4	4.6 ± 1.4	4.5 ± 1.1	1.70E-01	2.01E-01
VAS, mean ± SD	-	79.5 ± 10.0	21.5 ± 19.4	8.8 ± 22.1	0.0 ± 0.0	7.66E-31	9.96E-30
Arthralgia, No. (%)	-	61(100)	36(65.5)	6(20.7)	0(0)	9.45E-21	3.07E-20
JSS, mean ± SD	-	2.8 ± 0.8	0.6 ± 1.3	0.3 ± 0.8	0.0 ± 0.0	6.35E-28	2.75E-27
JTS, mean ± SD	-	3.2 ± 0.7	0.8 ± 0.9	0.4 ± 1.0	0.0 ± 0.0	3.19E-28	2.07E-27
Tophi, No. (%)	-	13(16.5)	12(16.7)	10(24.4)	0(0)	4.76E-01	5.16E-01

1 SD, standard deviation; BMI, body mass index; CRP, C-reactive protein; eGFR, estimated  
2 glomerular filtration rate; ESR, erythrocyte sedimentation rate; SCr, serum creatinine; SUA, serum  
3 uric acid; VAS, visual analog scale; JSS, joint swelling scores; JTS, joint tenderness scores.  
4 <sup>a)</sup> *P* value based on Kruskal-Wallis Test (continuous variables, or Wilcoxon rank-sum test for two  
5 groups) or Chi-Square (categorical variables, or Fisher's exact test for sparse counts) all five  
6 groups.