

Group	Subjects	Age	Male	Num. with 16S rRNA amplicon sequencing	Average reads	Num. with WGS sequencing	Average num. of reads
Gout	<i>n</i> = 35	50.31 ± 10.16	19/35	35/35	6107	16/35	59117661
Control	<i>n</i> = 33	48.73 ± 11.59	19/33	33/33	6594	18/33	59657047
Validation	<i>n</i> = 15	46.87 ± 11.46	10/15	15/15	6670	5/15	61574689
Group	BMI	Uric Acid*	Total Bilirubin*	GPT*	GOT*	ALP	Total Protein
Gout	23.09 ± 3.48	510.94 ± 91.47	14.92 ± 6.29	30.43 ± 9.75	24.31 ± 10.03	57.74 ± 11.81	67.97 ± 8.63
Control	22.93 ± 3.81	199.48 ± 67.14	9.16 ± 3.78	13.73 ± 5.18	11.55 ± 5.43	59.42 ± 14.18	69.79 ± 5.48
Validation	23.99 ± 3.47	402.73 ± 58.38	13.15 ± 4.67	22.93 ± 9.92	25.27 ± 6.38	67.40 ± 15.90	72.97 ± 8.47
Group	Urea Nitrogen*	Creatinine	Blood Glucose	Triglyceride	Cholesterol	HDLc	LDLc
Gout	7.49 ± 3.12	81.37 ± 20.61	5.73 ± 1.49	1.51 ± 0.64	4.61 ± 1.17	1.20 ± 0.19	2.69 ± 0.96
Control	4.53 ± 1.22	82.54 ± 13.71	5.12 ± 0.74	1.45 ± 0.59	4.74 ± 1.01	1.17 ± 0.28	2.67 ± 0.74
Validation	6.37 ± 4.36	78.78 ± 12.19	5.59 ± 1.01	1.35 ± 0.28	4.72 ± 0.85	1.20 ± 0.21	2.82 ± 0.51

Table 1. Sample information and blood index. NOTE: GPT represent glutamic-pyruvic transaminase; GOT represent glutamic-oxalacetic transaminase; HDLc represent high-density lipoprotein cholesterol; LDLc represent low-density lipoprotein cholesterol.

responsible enzyme xanthine dehydrogenase can be secreted by the *Escherichia coli* group of human intestinal bacteria^{16,17}. In uric acid catabolism, uricase, allantoinase and allantoinase activities can sequentially degrade uric acid to 5-hydroxyisourate, allantoin, allantoinate and eventually to urea. Synthesis of these enzymes was found vigorous in *Lactobacillus* and *Pseudomonas*, both common members of human intestinal microbiota¹⁸. Moreover transport protein of uric acid were found secreted by various indigenous microbes in human gut¹⁹. Therefore, we hypothesize that the intestinal microbiota can potentially serve as proxy to probe uric acid metabolism in the host for the purpose of diagnosis or prognosis.

To test this hypothesis, here we designed a cross-sectional study in an 83-member Chinese cohort that consists of both gout patients and healthy individuals. The taxonomic structure of their intestinal microbiota was determined by 16S rRNA gene pyrosequencing, and the functional profile of the corresponding microbiome was determined by shotgun metagenomic sequencing. Profound difference in microbial taxonomic and functional features was discovered between gout patients and healthy individuals. A microbiota-based model based on such difference was established and shown to diagnose gout at 88.9% accuracy.

Results

Intestinal microbiota altered profoundly in gout patients. The study cohort includes 83 Chinese adults (Table 1, Supplementary Table 1). The gout group consists of 35 adult patients who were clinically diagnosed by endocrinologists as gout based on blood uric acids, claim of joint pains and other parameters. The healthy group which serves as control consists of 33 healthy adults. An additional, 15-member group for model validation purpose consists of 6 gout patients and 9 healthy individuals. For each of the collectively 83 subjects, organismal structure of intestinal microbiota was analyzed by sequencing 16S rRNA gene amplicons, which revealed for each microbiota on average 202 operational taxonomic units (OTUs) from averagely 6402 reads (Supplementary Table 2). To characterize the functional profiles of microbiota, 16 from the gout group, 18 from the healthy group and 5 from the validation group were randomly selected for whole-metagenome shotgun sequencing, yielding 371.2 Giga base (Gb) of pair-end reads (averagely 59,681,612 high-quality reads for each microbiota; Supplementary Table 3).

Between the gout group and the healthy group, no significant difference was found for the factors of age, gender or BMI. However, among the array of blood indices, significant differences ($P < 0.001$, Wilcoxon rank-sum test) were detected between the two groups in blood uric acid value, total bilirubin, glutamic-pyruvic transaminase (GPT), glutamic oxalacetic transaminase (GOT) and urea nitrogen level (Fig. 1A; Supplementary Table 1). Intriguingly, in the gout group, the blood uric acid values of patients 14, 27, 39 and 60 were lower than other patients, which is consistent with the notion of insufficient sensitivity of the blood uric acid index for diagnosing gout in certain patients⁸.

To test whether any difference in organismal structure of intestinal microbiota was present, Principal Coordinates Analysis (PCoA) was performed based on the weighted Unifrac distances of 16S rRNA sequence profiles at the genus level. A significant reduction was observed in α diversity of intestinal microbiota from the gout group as compared to those from the healthy group ($P < 0.01$, Wilcoxon rank-sum test; Supplementary Table 4), suggesting that a lower microbial diversity in the intestine was associated with gout. Moreover, intestinal microbiota from the gout group and the healthy group were highly distinct in organismal structure, with the subjects forming two clusters that respectively corresponded to the two groups (Fig. 1B, upper panel), and significant separation ($P < 0.001$) in PC 1 was observed (Fig. 1B, lower panel; Wilcoxon rank-sum tests).

To correct for any possible population stratification caused by the non-gout-related factors such as age, gender and BMI, organismal structures of the microbiota were further analysed using Permutational Multivariate Analysis of Variance (PERMANOVA). After the correction those effects correlated with the non-gout-related factors disappeared (Supplementary Table 5). This further confirmed that the gout disease is one significant factor in explaining the observed variation in organismal structure of intestinal microbiota.

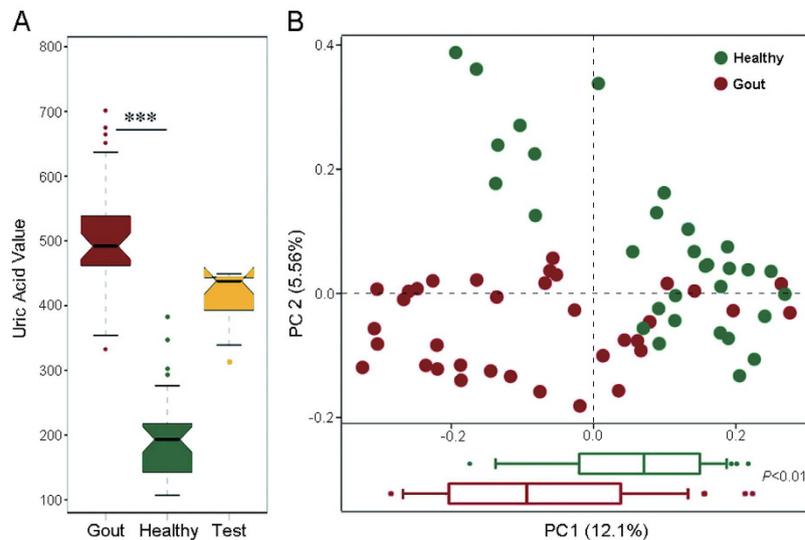


Figure 1. The composition of gut microbiota alters profoundly in gout patients. (A) The uric acid values in the gout patients, healthy (control) and validation groups. **(B)** A principal component (PCoA) score plot based on weighted UniFrac metrics for all participants. Each point represents the composition of the intestinal microbiota of one participant.

Furthermore, to probe the difference in functional profile of intestinal microbiota between gout patients and healthy subjects, high-quality reads from all samples were assembled and annotated for protein-coding genes, based on which a collective, non-redundant intestinal microbiota gene catalog for gout was created. Then for each of the sample, its reads were mapped to the collective gene catalog to reconstruct sample-specific gene profile, as well as the profiles for the associated Clusters of Orthologous Groups (COG, Supplementary Table 6) and Kyoto Encyclopedia of Genes and Genomes database Orthologues (KO; Supplementary Table 7). For the gene profiles, PCA analysis revealed that two clusters that respectively correspond to the gout group and the healthy group were formed (Supplementary Fig. 1A), and the separation in PCA 1 between the two groups was significant ($P < 0.001$; pairwise Wilcoxon tests). These results suggested a distinct functional gene structure of the gout patients as compared to that of the healthy subjects.

To test the degree of consistency between the observed clustering patterns for organismal structure and functional structure respectively, Procrustes analysis was performed based on the PCA matrix of genus-level organismal profile and that of functional gene profile (Supplementary Fig. 1B). The results revealed a strong correspondence between microbial taxonomy and function ($P < 0.001$, using 10,000 Monte Carlo label permutations). Thus the classification of the samples along the line of the gout group of subjects and the healthy group is robust, and moreover, both organismal structure and functional gene structure of intestinal microbiota from an individual host carry significant and consistent information for the gout disease state.

Microbial index of Gout. To explore microbial species associated with gout disease, microbial genes predicted from the whole-metagenome sequencing data were grouped into clusters based on the gene abundance profiles, with each of the clusters denoted a metagenomic species (MGS)²⁰. Among the 41 MGS identified from metagenomic species profile, 22 were enriched in healthy individuals, such as those MGS representing *Faecalibacterium prausnitzii*, *Clostridium* butyrate-producing bacterium and *Bifidobacterium pseudocatenulatum* (Wilcoxon rank-sum test; Fig. 2; Supplementary Table 8). On the other hand, 19 MGS were enriched in the gout patients, including species such as *Bacteroides caccae* and *Bacteroides xylanisolvens* (Fig. 2; Supplementary Table 8).

The highly distinct organismal features of gout suggested the possibility of classifying gout subjects based on the intestinal microbiota. On the other hand, based on the bacterial 16S rRNA pyrosequencing data, bacterial genera that were differentially distributed between the gout group and the healthy group were revealed. A total of 17 genera ($P < 0.05$, Wilcoxon rank-sum test) were associated with gout disease, with *Bacteroides*, *Holdemania*, *Anaerotruncus* and several others positively associated with the gout, whereas *Faecalibacterium*, *Coproccoccus*, *Ruminococcus* etc negatively associated with the gout (i.e., they were enriched in the healthy group; Table 2). These results were consistent with the organismal features identified via the MGS analysis above.

To test this hypothesis of microbiota-based diagnosis of gout, the 68-host cohort that includes the 33-member gout group and the 35-member healthy group was used as a training set for model construction. A ‘microbial index of gout’ (MiG) was derived based on the relative abundance of the 17 bacterial markers that distinguish between the healthy and the gout patients in the 68-host cohort:

$$\text{MiG} = \left(\frac{\sum_{i=9} \text{abundance}(\text{ggout enriched})_i}{9} \right) - \frac{\sum_{j=8} \text{abundance}(\text{ghealth enriched})_j}{8} \times 10^6 \quad (1)$$

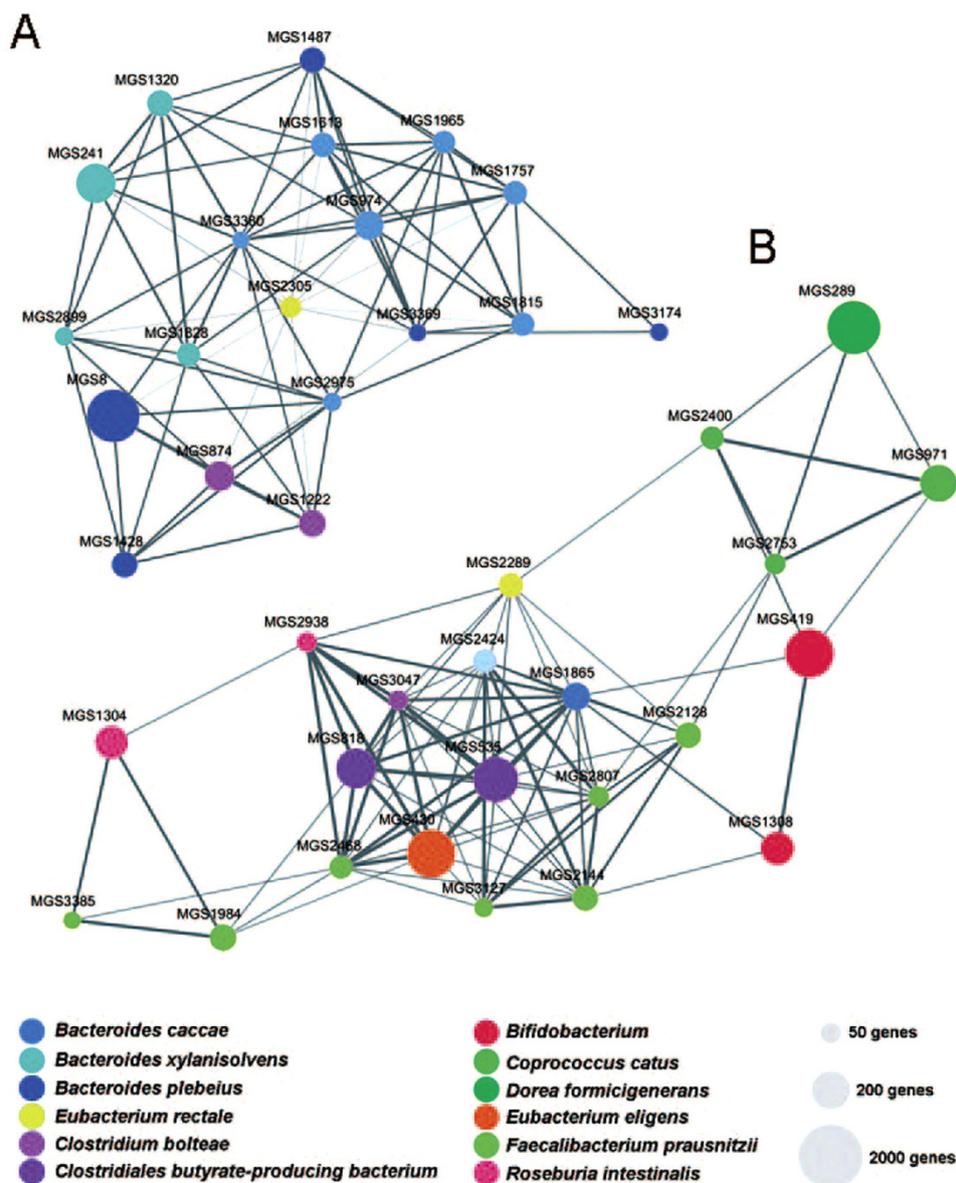


Figure 2. Taxonomic characterization of the intestinal microbiota in gout. Differentially abundant MGS networks enriched in gout patients ($n = 19$, panel A) and healthy individuals ($n = 22$, panel B). The edge width is proportional to the correlation strength. The node size is proportional to the mean abundance in the respective population. Nodes with the same color are classified in the same phylogenetic species. Every node represented a MGS.

Genus	P values (adjusted)	Odds ratios	Enriched	Genus	P values (adjusted)	Odds ratios	Enriched
<i>Coprococcus</i>	0.0005231	2.19	Control	<i>Barnesiella</i>	0.0041493	0.22	Gout
<i>Faecalibacterium</i>	0.0022616	9.62	Control	<i>Parasporobacterium</i>	0.0054789	0.06	Gout
<i>Oscillibacter</i>	0.0024974	1.00	Control	<i>Paraprevotella</i>	0.0093753	0.05	Gout
<i>Ruminococcus</i>	0.0064096	1.28	Control	<i>Anaerotruncus</i>	0.0106759	0.05	Gout
<i>Odoribacter</i>	0.0115741	0.17	Control	<i>Pseudobutyrvibrio</i>	0.0319377	0.19	Gout
<i>Subdoligranulum</i>	0.0210841	2.56	Control	<i>Bacteroides</i>	0.0368476	49.67	Gout
<i>Robinsoniella</i>	0.0286258	0.01	Control	<i>Holdemanina</i>	0.0431658	0.02	Gout
<i>Dialister</i>	0.0325178	0.38	Control	<i>Acetanaerobacterium</i>	0.047967	0.01	Gout
<i>Alistipes</i>	0.0360435	0.77	Control				

Table 2. Microbial biomarkers of the gout disease.

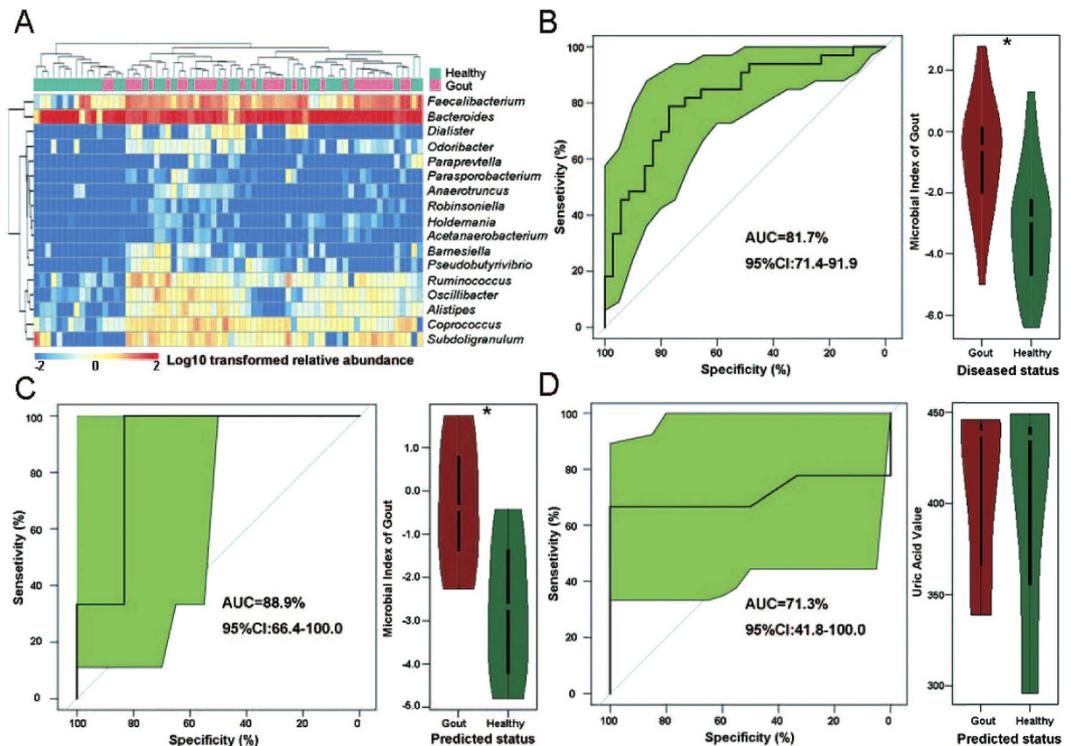


Figure 3. Classification of the gout status using bacterial genus-level biomarkers based on 16S pyrosequencing data and stratification of RISK hosts in a validation cohort. (A) The heatmap indicated the ability of the genus-level biomarkers to discriminate the healthy and gout groups. (B) Accuracy of the microbiota-based predictive model is measured by AUC in the gout and the control groups, and the box figure of MiG for all participants in the gout and the control groups were shown. (C) Accuracy of the microbiota-based model is measured by AUC in the validation group. (D) Accuracy of blood uric acid value based model is measured by AUC in the validation group.

For evaluation of gout disease based on these biomarkers, the correlation between the biomarkers and subjects were shown as a heat map, which indicated the ability of these biomarkers to discriminate healthy and gout hosts (Fig. 3A). Moreover, the Wilcoxon rank-sum test revealed that the MiGs between the patient group and the control group were significant ($P < 0.05$; Fig. 3B). Furthermore, performance of the MiG model in the training data (the 68-member cohort) was assessed by receiver operating characteristic curve (ROC). From the ROC of training dataset, we have determined the Youden index and the threshold of MiG as -2.157 . For a given subject, a MiG that reaches this threshold would suggest increased risk of suffering gout. The area under the receiver operating characteristic (ROC) curves was 81.7% (Fig. 3B), indicating the microbial gout index could be used to classify gout individuals at high accuracy.

To validate the performance of this microbial model, an additional 15 subjects were included as the validation group. The blood uric acid values of these 15 subjects were higher than those of the control group but lower than that of severe gout patients (Supplementary Table 1). Thus uric acid value alone was not able to reliably diagnose gout for these individuals. Subsequent return-visits to clinics (six months from the initial diagnosis) confirmed that 6 subjects in this group suffered gout disease and 9 individuals were healthy. Analysis of their intestinal microbiota revealed that the accuracy of diagnosis via the microbial index based predictive model in the validation group reached 88.9% (Fig. 3C), higher than that based on blood uric acid (71.3%; Fig. 3D).

To probe the potential link between the organismal biomarkers of gout and the various blood indexes, Spearman rank correlations between the abundance matrixes of microbial biomarkers and each of the blood indexes was performed (Supplementary Fig. 3A). Interestingly, positive correlation was observed between the amounts of uric acid, bilirubin, glutamic-pyruvic transaminase (GPT) and glutamic-oxalacetic transaminase (GOT) in the blood and the genera enriched in gout patients (such as *Bacteroides*, *Holdemania*, *Anaerotruncus*, etc), revealing a number of potential bacteria taxa whose activities might be implicated in the development of gout (Supplementary Fig. 3A).

Mechanisms underlying Microbial Index of Gout. To probe the potential mechanism underlying the Microbial Indices of Gout, the shotgun metagenomic data were analyzed at the functional and metabolic pathway level. The results revealed that 5,245 KOs and 2,286 COGs were either positively or negatively associated with gout. The gout-associated KOs were then mapped to the reference metabolism pathway, which revealed that the gout patients were enriched in the metabolism of purine, starch and sucrose, sphingolipid, alanine, aspartate and glutamate, retinol, porphyrin and chlorophyll (Supplementary Table 9). On the pathway of purine metabolism,

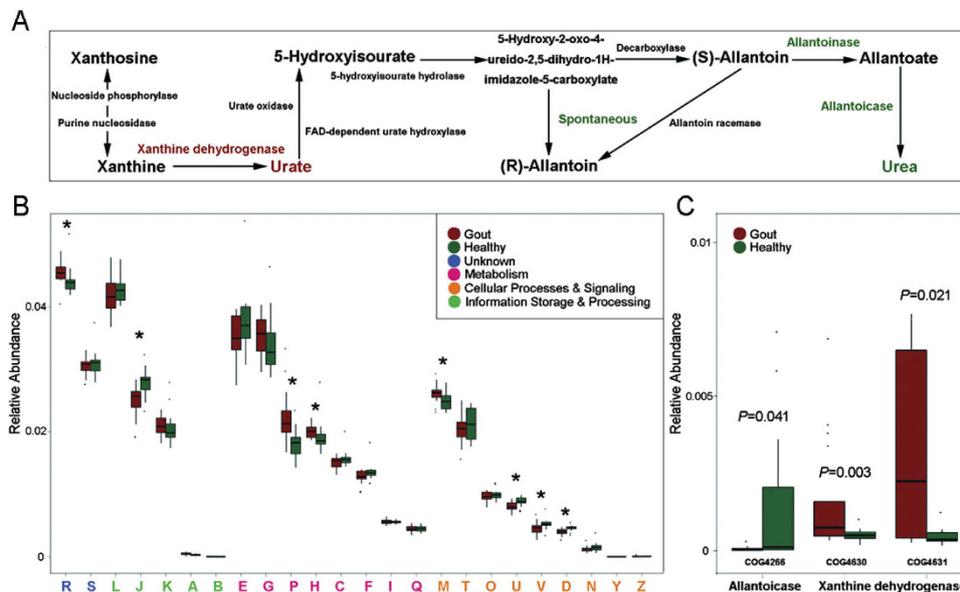


Figure 4. Functional features of gut microbiota in gout. (A) The metabolism of purine degradation. The enzymes in red were enriched in the gout patient group and those in green were enriched in the healthy (control) group. (B) Comparison of COGs between the patient and the control groups. A: RNA processing and modification; J: Translation, ribosomal structure and biogenesis; D: Cell cycle control, cell division, chromosome partitioning; M: Cell wall/membrane/envelope biogenesis; U: Intracellular trafficking, secretion, and vesicular transport; V: Defence mechanisms; H: Coenzyme transport and metabolism; P: Inorganic ion transport and metabolism; R: General function prediction only. The capital letters in red represent those functions enriched in the gout patient group, while the capital letters in blue represent those functions enriched in the control group. (C) Comparison of distribution of COGs of xanthine dehydrogenase and allantoicase between the patient and the control groups.

the xanthine dehydrogenase which can degrade the purine to uric acid was enriched in the gout patients, whereas the allantoicase that degrades the uric acid to urea was depleted (Fig. 4A). Thus it is possible that in the intestinal microbiota of gout patients, a significant amount of purine were degraded to uric acid which however could not be further degraded to urea, leading to abnormal accumulation of uric acid in gout patients. In contrast, the KO enriched in the healthy group included those involved in butyric acid biosynthesis, inositol phosphate metabolism, propanoate metabolism, methane metabolism, glycerolipid metabolism, thiamine metabolism and nitrolo- uene degradation (Supplementary Table 9).

Analysis of the functional feature as defined by COGs revealed that the xanthine dehydrogenase (COG4360 and COG4361) was significantly enriched ($P < 0.05$, Wilcoxon rank-sum test, Fig. 4C) in gout patients, but the allantoicase (COG4266) depleted. These results were consistent with those based on the metabolic pathway of purine degradation. Moreover, the COGs that were enriched in gout patients include those implicated in RNA processing and modification, Cell wall/membrane/envelope biogenesis, Coenzyme transport and metabolism and Inorganic ion transport and metabolism (Fig. 4B and Supplementary Table 10). In contrast, COGs of Translation, ribosomal structure and biogenesis, Cell cycle control/division, Intracellular trafficking/secretion/vesicular transport and Defence mechanisms were enriched in the healthy group.

Discussion

In current clinical practice, the blood uric acid value is considered to be the reference index in clinical gout diagnosis, however it can be hysteretic. Our diagnosis model for gout based on metagenomic genera of intestinal microbiota sampled from stool has shown promise in improving the sensitivity of diagnosis, where the model based on 17 genera achieved 88.9% accuracy. In fact, patients in the validation group that could not be reliably diagnosed by blood uric acid value can be correctly diagnosed as gout using the microbial index. Therefore the MiG may potentially be a valuable tool for early detection of gout.

The MiG revealed the presence of bacterial dysbiosis in gout patients. In the network of MGS, *Bacteroides caccae* ($n = 7$) and *Bacteroides xylanisolvens* ($n = 4$) were enriched in gout patients. In past studies, *Bacteroides caccae* was recognized one biomarker of IBD, and the *OmpW* protein produced by *B. caccae* was a target of the IBD-associated immune response²¹, thus the enriched intestinal *B. caccae* in gout patients could potentially induce serious inflammatory response. *Bacteroides xylanisolvens* strains isolated from feces of human and other animals was considered to have no virulence and be safe for food use^{22,23}. Significant depletion of *Faecalibacterium prausnitzii* which was reported to have anti-inflammatory properties and contribute to gut health through butyrate production²⁴ was observed for gout patients, potentially explaining the decline in butyric acid biosynthesis in gout patients. Protective mechanisms of butyric acid in human intestine include providing the nutrition for intestinal mucosa, promoting the growth and repairing of intestinal villus, enhancing the intestinal immunity,

	Metagenomic Species	T2D	Gout	LC		Metagenomic Species	T2D	Gout	LC
Enriched species in T2D/ Gout/LC	<i>Akkermansia muciniphila</i>	+	n/a	n/a	Depleted species in T2D/ Gout/LC	<i>Faecalibacterium prausnitzii</i>	–	–	–
	<i>Bacteroides intestinalis</i>	+	n/a	n/a		<i>Haemophilus parainfluenzae</i>	–	n/a	n/a
	<i>Clostridium bolteae</i>	+	+	n/a		<i>Roseburia intestinalis</i>	–	–	n/a
	<i>Clostridium hathewayi</i>	+	n/a	n/a		<i>Roseburia inulinivorans</i>	–	n/a	n/a
	<i>Clostridium ramosum</i>	+	n/a	n/a		<i>Eubacterium rectale</i>	–	–	n/a
	<i>Clostridium symbiosum</i>	+	n/a	n/a		<i>Bacteroides stercoris</i>	n/a	–	n/a
	<i>Eggerthella lenta</i>	+	n/a	n/a		<i>Bifidobacterium dentium</i>	n/a	–	n/a
	<i>Escherichia coli</i>	+	n/a	n/a		<i>Bifidobacterium pseudocatenulatum</i>	n/a	–	n/a
	<i>Bacteroides caccae</i>	n/a	+	n/a		<i>Clostridium bolteae</i>	n/a	–	n/a
	<i>Bacteroides xylanisolvens</i>	n/a	+	n/a		<i>Coprococcus catus</i>	n/a	–	n/a
	<i>Bacteroides ovatus</i>	n/a	+	n/a		<i>Dorea formicigenerans</i>	n/a	–	n/a
	<i>Eubacterium rectale</i>	n/a	+	n/a		<i>Bacteroides plebeius</i>	n/a	–	n/a
	<i>Lactobacillus fermentum</i>	n/a	n/a	+		<i>Bacteroides clarus</i>	n/a	n/a	–
	<i>Lactobacillus mucosae</i>	n/a	n/a	+		<i>Bacteroides uniformis</i>	n/a	n/a	–
	<i>Lactobacillus salivarius</i>	n/a	n/a	+		<i>Bilophila wadsworthia</i>	n/a	n/a	–
	<i>Veillonella atypica</i>	n/a	n/a	+		<i>Coprococcus comes</i>	n/a	n/a	–
	<i>Veillonella dispar</i>	n/a	n/a	+		Metabolic Pathway	T2D	Gout	LC
	<i>Veillonella parvula</i>	n/a	n/a	+		Butyrate biosynthesis	–	–	–
	<i>Megasphaera micronuciformis</i>	n/a	n/a	+		Cell motility	–	n/a	n/a
	<i>Prevotella buccae</i>	n/a	n/a	+		Cofactors & vitamins Metabolism	–	n/a	n/a
<i>Clostridium perfringens</i>	n/a	n/a	+	CH4 metabolism	n/a	–	n/a		
<i>Ruminococcus gnavus</i>	n/a	n/a	+	Amino acid metabolism	n/a	n/a	–		
<i>Streptococcus anginosus</i>	n/a	n/a	+	Carbohydrate metabolism	n/a	n/a	–		
<i>Streptococcus oralis</i>	n/a	n/a	+	Energy metabolism	n/a	n/a	–		
<i>Streptococcus salivarius</i>	n/a	n/a	+	Signal transduction	n/a	n/a	–		

Table 3. Comparison of taxonomy and functional features in different chronic diseases.

facilitating the growth of beneficial microbes and inhibiting the pathogenic bacteria²⁵, thus decline in butyric acid biosynthesis can cause a number of physiological disorders in the host. On the other hand, the microbial xanthine dehydrogenase which can degrade the purine to uric acid was enriched in the gout patients whereas the microbial allantoinase that degrades the uric acid to the urea was depleted in gut patients. Thus, another possible link between activity of intestinal microbiota and gout pathogenesis might be that the overly abundant xanthine dehydrogenase and the relative deficiency of allantoinase in the intestinal microbiota might have accumulated more uric acid and consequently aggravated the gout symptoms.

Microbial dysbiosis in intestinal microbiota for two other chronic metabolic disorders, liver cirrhosis²⁶ and T2D²⁷, were recently reported, also based on cohorts of Chinese adults. We found that a decreased level in microbial metabolism of butyrate biosynthesis was a shared feature of the three kinds of diseases^{26,27}. A remarkable similarity of reduction of *Faecalibacterium prausnitzii* and inhibition of butyrate biosynthesis was observed in patients from all the three chronic diseases (Table 3). This observation highlights the potential of *Faecalibacterium prausnitzii* as an indicator of intestine health.

However, in terms of the organismal structure of intestinal microbiota, T2D and gout patients were more similar, while both quite distinct from that of liver cirrhosis (Supplementary Fig. 4), in that patients suffering from T2D and gout diseases exhibit a lack of the species *Faecalibacterium prausnitzii*, *Roseburia intestinalis* and *Eubacterium rectale* (when compared with healthy individuals). Moreover, certain host features affected by intestinal microbiota similar to that observed for gout patients was also present in patients of T2D (Table 3 and Supplementary Fig. 4)^{27,28}. In T2D patients, blood uric acid value are usually elevated. In fact, the diabetogenic effect of uric acid has been reported since 1950²⁹, and here a positive correlation between blood uric acid and glucose values was also observed (Supplementary Fig. 3B). Moreover, many believed T2D as one serious complication of gout, and even considered the uric acid value as one indirect index for T2D diagnosis³⁰. However, the etiological link between uric acid and T2D is not yet fully understood, although it was reported that uric acid itself might contribute to the exacerbation of insulin resistance³¹. Our results here suggested that a similar or shared functional disorder of intestinal microbiota might underlie the pathological link between the two metabolic syndromes.

Catalogs of reference genes in the human gut microbiome are crucial for functional metagenomic analyses in healthy and diseased populations. A large catalog of reference genes from intestinal microbiota that included 9,879,896 genes was recently reported³², however it was mainly derived from healthy subjects from MetaHIT and HMP, which might limit its reference value in help interpreting diseased microbiota. This point was supported by a recent study linking dysbiosis of intestinal microbiota to liver cirrhosis which concluded that the number of unique genes among gene sets of different disease, genotypes, ages and dietary habits can be large²⁶. Thus the intestinal gene catalogue of gout established in this study that includes 2,653,431 microbial genes from both gout

and healthy cohorts should be of value for future studies probing the role of intestinal microbiota in gout and related disorders.

The strong association between individual bacterial taxa and gout revealed here suggested the importance of including intestinal microbiota metabolism into the diagnosis and mechanistic interrogation of gout. For example, examination of the link between microbiota activity with host genetic disorders, such as dysfunction of human ABCG2 (a high-capacity urate transporter regulating serum uric acid levels, which was identified as one major reason for early-onset gout³³), might reveal the interplay between microbiota and hosts in disease initiation and development. Moreover, tracking the structure and functional activity of intestinal microbiota in longitudinal studies in host populations of large size or from distinct ethnic background or diet pattern should reveal the degree and nature of heterogeneity in microbiota among host individuals and promise to provide new ways of gout prevention and control.

Materials and Methods

Study design and subject recruitment. The study cohort was composed of three groups. Group one (case group) consisted of 35 gout patients, aged 32–75 years, and they were recruited from the First Affiliated Hospital of Baotou Medical College, located in Baotou, Inner Mongolia Autonomous Region of China. Diagnosis of gout was confirmed by the analysis of blood uric acid among patients claiming of painful joints. The second group (control group) consisted of 33 healthy individuals, aged 28–70 years. The validation group consisted of 15 subjects that were aged 28–69 years, among whom were 6 gout patients and 9 healthy individuals. When recruiting subjects for test group, we chose individuals with blood uric acid value higher than control group level but lower than serious gout patients' level. All the subjects were asked to fill up a dietary diary, which recorded information such as their gender, age and diet for a period of three consecutive days prior to the collection of fecal samples (Supplementary Table 1). The study was approved by the Ethical Committee of the First Affiliated Hospital of Baotou Medical College (Baotou, China), and informed consent was obtained from all the 83 volunteers before enrollment in the study. Sampling and all subsequent steps described in the Materials and Methods have been conducted in accordance with the approved guidelines.

Fecal samples were collected from each individual in the morning before the first meal. After the weight of the fecal sample was determined, RNAlater[®] solution (Ambion, USA) was added into the tube in the ratio of one part fecal sample to five parts RNAlater[®] solution and mixed homogeneously via vortexing and stored at -20°C until further processing. All of the 83 samples were used for bacterial 16S rRNA genes V1-V3 region pyrosequencing and 39 samples (include 17 gout patients, 18 control individuals and 5 test individuals) were selected for whole-genome shotgun (WGS) sequencing.

DNA extraction. The QIAamp[®] DNA Stool Mini Kit (Qiagen, Hilden, Germany) was used for DNA extraction from the fecal samples. The quality of the extracted DNA was assessed by 0.8% agarose gel electrophoresis, and the OD 260/280 was measured by spectrophotometry. All of the DNA samples were stored at -20°C until further processing.

PCR amplification of the bacterial 16S rRNA genes V1-V3 region and pyrosequencing. The V1-V3 region of the 16S ribosomal RNA (rRNA) genes was amplified for barcoded pyrosequencing for all 83 volunteers. A set of 10-nucleotide barcodes was added to the universal forward primer 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and the reverse primer 533R (5'-TTACCGCGGCTGCTGGCAC-3'), which were targeted at the domains Bacteria. PCR amplification was then performed as described previously³⁴.

Quality of the PCR products was ensured using Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, Calif.) in accordance with the manufacturer's instructions. The PCR products were pooled in equimolar ratios with a final concentration of 100 nmol/L for pyrosequencing (Roche GS FLX) performed by Shanghai Majorbio Bio-pharm Technology Co., Ltd.

Pyrosequencing sequence processing and bioinformatics analysis. High-quality sequences were extracted from the raw reads and the extracted sequences were sorted into different samples according to the barcodes. After the removal of the barcodes and primer, the extracted sequences were processed mainly using the QIIME (v1.5.0) suite of software tools³⁵.

By pyrosequencing, we generated a dataset consisting of 535,153 high-quality 16S rRNA gene sequences with an average of 6,398 sequences (range = 3,817–9,983, SD = 1,433) obtained for each of the 83 samples (Supplementary Table 2). Operational taxonomic units (OTUs) were delineated at a 97% similarity level, leaving 3,685 for further analysis.

Whole-genome shotgun (WGS) sequencing and quality control. All samples were sequenced in the Illumina HiSeq2000 instrument. Libraries were prepared with a fragment length of approximately 300 bp. Paired-end reads were generated with 100 bp in the forward and reverse directions. The length of each read was trimmed with Sickle. Reads that aligned to the human genome were also removed. This set of high-quality reads was then used for further analysis. An average of 9.52 gigabases (Gb) paired-end reads were obtained for each sample, totaling 371.2 Gb of high-quality data that were free of human DNA and adaptor contaminants (Supplementary Table 3).

Illumina short reads *de novo* assembly, gene prediction and construction of the non-redundant gene catalogue. The Illumina reads were assembled into contigs using IDBA-UD³⁶ with default parameters. Genes were predicted on the contigs with MetaGeneMark³⁷. A non-redundant gene catalogue was constructed with CD-HIT³⁸ using a sequence identity cut-off of 0.95, with a minimum coverage cut-off of 0.9 for the shorter sequences. This catalogue contained 2,653,431 microbial genes.

Functional annotation. We aligned putative amino acid sequences, which were translated from the gene catalogue, against the proteins/domains in COG and KEGG databases using blastp (e-value $\leq 1e-5$ with a bit-score higher than 60). Each protein was assigned to the KEGG orthologue group (KO) or cluster of orthologous group (COG) by the highest scoring annotated hit.

Computation of relative gene abundance. To assess the abundance of genes, reads were aligned to the gene catalogue with Bowtie2³⁹ using parameters: -p 12 -x nt -1 R1.fastq -2 R2.fastq -S R.sam. Then, for any sample N, we calculated the abundance as follows:

Step 1: Calculation of the copy number of each gene:

$$b_i = \frac{x_i}{L_i} \quad (2)$$

Step 2: Calculation of the relative abundance of gene i

$$a_i = \frac{b_i}{\sum_i b_i} \quad (3)$$

a_i : the relative abundance of gene i

b_i : the copy number of gene i from sample N

L_i : the length of gene i

x_i : the number of mapped reads

Metagenomic species (MGS). The co-abundance principle and canopy clustering algorithm were applied to generate the metagenomic species (MGS). The detailed approach was described in previous references^{20,26}. MGS were assigned to a given genome when more than 80% of its sub-gene matched the same genome using blastN, at a threshold of 95% identity over 90% of gene length. Those MGS that cannot meet the above conditions were annotated using blastp analysis. If over 80% of the genes from a MGS had the same taxonomical level of assignment, this MGS was identified as the organism.

Statistical analysis. All statistical analyses were made in the R software. We performed the PERMANOVA⁴⁰ (Permutational multivariate analysis of variance) analysis using the method implemented in R package “vegan”, and the permuted P -value was obtained by 10,000 times permutations. PCA and Procuste analysis was performed in R using the ade4⁴¹ package. Differential abundance of genus, genes and COGs were tested by Wilcoxon rank sum test, and the significant different ($p < 0.01$) genus, genes and COGs were considered as the potential gout biomarkers. We applied the Receiver Operator Characteristic (ROC) analysis to assess the performance of the gout metagenomic biomarkers by using the “pROC” package in R software⁴². The correlation between biomarkers and blood index were calculated by Spearman rank correlation coefficient and visualized by heatmap in R using the “pheatmap” package.

References

- Burns, C. M. & Wortmann, R. L. Gout therapeutics: new drugs for an old disease. *Lancet* **377**, 165–177 (2011).
- Richette, P. & Bardin, T. Gout. *Lancet* **375**, 318–328 (2010).
- Zhu, Y., Pandya, B. J. & Choi, H. K. Comorbidities of gout and hyperuricemia in the US general population: NHANES 2007–2008. *Am J Med* **125**, 679–687 (2012).
- Kuo, C. F., Grainge, M. J., Mallen, C., Zhang, W. & Doherty, M. Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study. *Ann Rheum Dis* **74**, 661–667 (2014).
- Guo, M. *et al.* Polymorphism of rs7688672 and rs10033237 in cGKI/PRKG2 and gout susceptibility of Han population in northern China. *Gene* **562**, 50–54 (2015).
- Krishnan, E. *et al.* Serum urate and incidence of kidney disease among veterans with gout. *J Rheumatol* **40**, 1166–1172 (2013).
- Lee, S. *et al.* Uric Acid Level Is Not An Independent Predictor Of Cardiovascular Diseases In Gout Patients With Treatment; Long-Term Follow-Up Data In Single Tertiary Center In South Korea. *Arthritis and rheumatism* **65**, S856–S856 (2013).
- McQueen, F. M. Gout in 2013. Imaging, genetics and therapy: gout research continues apace. *Nat Rev Rheumatol* **10**, 67–69 (2014).
- Chang, W. C. Dietary intake and the risk of hyperuricemia, gout and chronic kidney disease in elderly Taiwanese men. *Aging Male* **14**, 195–202 (2011).
- Sorensen, L. B. & Levinson, D. J. Origin and extrarenal elimination of uric acid in man. *Nephron* **14**, 7–20 (1975).
- Nicholson, J. K. *et al.* Host-gut microbiota metabolic interactions. *Science* **336**, 1262–1267 (2012).
- Kau, A. L., Ahern, P. P., Griffin, N. W., Goodman, A. L. & Gordon, J. I. Human nutrition, the gut microbiome and the immune system. *Nature* **474**, 327–336 (2011).
- Jumpertz, R. *et al.* Energy-balance studies reveal associations between gut microbes, caloric load, and nutrient absorption in humans. *Am J Clin Nutr* **94**, 58–65 (2011).
- Gordon, J. I., Dewey, K. G., Mills, D. A. & Medzhitov, R. M. The Human Gut Microbiota and Undernutrition. *Sci Transl Med* **4**, (2012).
- Xu, J. & Gordon, J. I. Honor thy symbionts. *Proc Natl Acad Sci USA* **100**, 10452–10459 (2003).
- Crane, J. K., Naeher, T. M., Broome, J. E. & Boedeker, E. C. Role of host xanthine oxidase in infection due to enteropathogenic and Shiga-toxicogenic *Escherichia coli*. *Infect Immun* **81**, 1129–1139 (2013).
- Sathisha, K. R. *et al.* Synthesis and xanthine oxidase inhibitory activity of 7-methyl-2-(phenoxyethyl)-5H-[1,3,4]thiadiazolo[3,2-a]pyrimidin-5-one derivatives. *Bioorg Med Chem* **19**, 211–220 (2011).
- Hsieh, C. Y., Lin, H. J., Chen, C. H., Lai, E. C. & Yang, Y. H. Chronic kidney disease and stroke. *Lancet Neurol* **13**, 1071 (2014).
- Hosomi, A., Nakanishi, T., Fujita, T. & Tamai, I. Extra-renal elimination of uric acid via intestinal efflux transporter BCRP/ABCG2. *PLoS One* **7**, e30456 (2012).
- 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).
- Wei, B. *et al.* Molecular cloning of a *Bacteroides caccae* TonB-linked outer membrane protein identified by an inflammatory bowel disease marker antibody. *Infect Immun* **69**, 6044–6054 (2001).

22. Ulsemer, P. *et al.* Safety and tolerance of *Bacteroides xylanisolvens* DSM 23964 in healthy adults. *Benef Microbes* **3**, 99–111 (2012).
23. Ulsemer, P., Toutounian, K., Schmidt, J., Karsten, U. & Goletz, S. Preliminary safety evaluation of a new *Bacteroides xylanisolvens* isolate. *Appl Environ Microbiol* **78**, 528–535 (2012).
24. Eckburg, P. B. *et al.* Diversity of the human intestinal microbial flora. *Science* **308**, 1635–1638 (2005).
25. Li, G., Yao, W. & Jiang, H. Short-chain fatty acids enhance adipocyte differentiation in the stromal vascular fraction of porcine adipose tissue. *J Nutr* **144**, 1887–1895 (2014).
26. Qin, N. *et al.* Alterations of the human gut microbiome in liver cirrhosis. *Nature* **513**, 59–64 (2014).
27. Qin, J. J. *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55–60 (2012).
28. Karlsson, F. H. *et al.* Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* **498**, 99–103 (2013).
29. Griffiths, M. The mechanism of the diabetogenic action of uric acid. *J Biol Chem* **184**, 289–298 (1950).
30. Tsouli, S. G., Liberopoulos, E. N., Mikhailidis, D. P., Athyros, V. G. & Elisaf, M. S. Elevated serum uric acid levels in metabolic syndrome: an active component or an innocent bystander? *Metabolism* **55**, 1293–1301 (2006).
31. Facchini, F., Chen, Y. D., Hollenbeck, C. B. & Reaven, G. M. Relationship between resistance to insulin-mediated glucose uptake, urinary uric acid clearance, and plasma uric acid concentration. *Jama* **266**, 3008–3011 (1991).
32. Li, J. *et al.* An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol* **32**, 834–841 (2014).
33. Matsuo, H. *et al.* ABCG2 dysfunction causes hyperuricemia due to both renal urate underexcretion and renal urate overload. *Sci Rep* **4**, 3375 (2014).
34. Caporaso, J. G. *et al.* PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **26**, 266–267 (2010).
35. Claesson, M. J. *et al.* Comparative Analysis of Pyrosequencing and a Phylogenetic Microarray for Exploring Microbial Community Structures in the Human Distal Intestine. *PLoS One* **4**, e6669 (2009).
36. Peng, Y., Leung, H. C. M., Yiu, S. M. & Chin, F. Y. L. IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **28**, 1420–1428 (2012).
37. Zhu, W. H., Lomsadze, A. & Borodovsky, M. Ab initio gene identification in metagenomic sequences. *Nucleic Acids Research* **38**, e132 (2010).
38. Li, W. Z. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
39. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**, R25 (2009).
40. McArdle, B. H. & Anderson, M. J. Fitting multivariate models to community data: A comment on distance-based redundancy analysis. *Ecology* **82**, 290–297 (2001).
41. Zapala, M. A. & Schork, N. J. Multivariate regression analysis of distance matrices for testing associations between gene expression patterns and related variables. *Proc Natl Acad Sci USA* **103**, 19430–19435 (2006).
42. Robin, X. *et al.* pROC: an open-source package for R and S plus to analyze and compare ROC curves. *Bmc Bioinformatics* **12**, 77 (2011).

Acknowledgements

We sincerely thank all the volunteers for their participation.

Author Contributions

H.Z. and Y.K.L. designed the study. Z.W. and J.Z. collected clinical samples. Z.G., K.Y.A., J.Q., L.W., E.K. and H.D. processed and sequenced samples. J.Z., S.H., Q.H., Y.Z., X.S. and Z.G. analyzed data. J.Z., J.X. and S.H. wrote the paper.

Additional Information

Data Availability: The sequence data reported in this paper have been deposited in the MG-RAST database (Project No. 9870).

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Guo, Z. *et al.* Intestinal Microbiota Distinguish Gout Patients from Healthy Humans. *Sci. Rep.* **6**, 20602; doi: 10.1038/srep20602 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>