



OPEN

DATA DESCRIPTOR

A Comprehensive Dataset on Microbiome Dynamics in Rheumatoid Arthritis from a Large-Scale Cohort Study

Jing Li^{1,14}, Jun Xu^{2,3,14}, Jiayang Jin¹, Congmin Xu⁴, Yuzhou Gan¹, Yifan Wang¹, Ruiling Feng¹, Wenqiang Fan⁵, Yingni Li¹, Xiaozhen Zhao¹, Yucui Li⁶, Shushi Gong⁷, Linchong Su⁷, Yueming Cai⁸, Lianjie Shi⁹, Xiaolin Sun¹⁰, Yang Xiang⁷, Qingwen Wang⁸, Ru Li¹, Jinxia Zhao¹⁰, Yulan Liu^{2,3}, Junjie Qin¹¹, Zhanguo Li^{1,12,13}✉ & Jing He¹✉

Alterations in intestinal microbiota have been identified as a key risk factor in rheumatoid arthritis (RA). This study presents a multidimensional gut microbiota profile from a large cohort of RA patients, stratified by disease stage and treatment regimens, and compared to healthy controls. Our dataset comprises gut microbiota profiles from 2,238 individuals, including 1,034 RA patients (Asia Pacific RA cohort, APRAC) and 1,204 healthy controls. This dataset is enriched with detailed clinical metadata, including patient profiles, treatment histories, and environmental factors, providing a comprehensive “disease exposome” for RA. By integrating 16S rRNA gene sequencing with demographic, clinical, and environmental data, we offer a valuable resource to explore the complex relationships between gut microbiota and RA progression. This large-scale dataset is expected to be a foundation for collaborative research, advancing our understanding of the microbiome’s systemic effects in RA and other autoimmune diseases and potentially guiding new therapeutic approaches.

Background & Summary

Rheumatoid arthritis (RA) is a chronic, systemic autoimmune disease characterized by inflammation of the synovial joints, leading to joint destruction, disability, and decreased quality of life^{1–4}. While the etiology of RA involves a complex interplay of genetic and environmental factors, recent studies suggest that dysbiosis of the gut microbiota plays a crucial role in the disease’s pathogenesis and progression. The gut microbiota, consisting of trillions of microorganisms, has emerged as a critical player in human health, influencing metabolic, immunologic, and even psychological processes. Alterations in the composition and function of this microbial community have been associated with a variety of diseases, including autoimmune disorders like RA^{1,5,6}.

¹Department of Rheumatology and Immunology, Peking University People’s Hospital, Beijing Key Laboratory for Rheumatism Mechanism and Immune Diagnosis (BZ0135), Beijing, 100044, China. ²Department of Gastroenterology, Peking University People’s Hospital, Beijing, 100044, China. ³Clinical Center of Immune-Mediated Digestive Diseases, Peking University People’s Hospital, Beijing, 100044, China. ⁴BioMap (Beijing) Intelligent Technology Co., Ltd. (BioMap), Beijing, 100044, China. ⁵Department of Rheumatology and Immunology, Xinxiang Central Hospital, Henan, 45300, China. ⁶Third Hospital of Shanxi Medical University, Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Tongji Shanxi Hospital, Taiyuan, 030032, China. ⁷Department of Rheumatology and Immunology, Minda Hospital of Hubei Minzu University, Hubei, 445000, China. ⁸Department of Rheumatology and Immunology, Peking University Shenzhen Hospital, Shenzhen, Guangdong, 518000, China. ⁹Department of Rheumatology and Immunology, Peking University Shougang Hospital, Beijing, 100144, China. ¹⁰Department of Rheumatology and Immunology, Peking University Third Hospital, Beijing, 100191, China. ¹¹Promegene Institute, Shenzhen, 518110, China. ¹²State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing, 100191, China. ¹³Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, Beijing, 100091, China. ¹⁴These authors contributed equally: Jing Li, Jun Xu. ✉e-mail: li99@bjmu.edu.cn; hejing1105@126.com

Studies on the relationship between the gut microbiome and RA using next-generation sequencing have been conducted to explore how alterations in the gut microbiota may influence the development and progression of RA^{1,5}. However, the existing research has faced inconsistencies and controversies, mainly due to limited sample sizes, a lack of stratification by disease stage and treatment status, and insufficient depth of microbial profiling. Furthermore, the temporal dynamic changes in gut microbial communities related to RA onset, progression, and treatment remain poorly understood.

To address these knowledge gaps, we comprehensively analyzed the gut microbiota in a large, stratified cohort of 2,238 individuals, including 1,034 RA patients and 1,204 healthy controls. For the profiling of the bacterial communities, 16S ribosomal RNA (rRNA) V3-V4 amplicon sequencing was performed after fecal samples collection, DNA extraction, and library construction as described in the Methods section. The resulting sequencing dataset elucidated the role of gut microbiota in RA, spanning various stages of the disease, treatment regimens, and clinical outcomes.

The purpose of presenting our data is to communicate our research findings transparently, allowing for the validation of results and contributing to the collective knowledge in the field. We believe this approach leads to more comprehensive studies integrating diverse methodologies and perspectives, thus advancing our understanding of RA and its link with gut microbiota.

Methods

Study design. All participants were recruited from six centers, including Peking University People's Hospital (Beijing 1) and Peking University Third Hospital (Beijing 2), Xinxiang Central Hospital in Henan, Shanxi Grand Hospital in Shanxi, Enshi Huiyi Hospital of Rheumatic Diseases in Hubei, and Peking University Shenzhen Hospital in Guangdong (Fig. 1A). It is important to note that certain facilities, such as those in Hubei and Beijing2, contributed only RA data and not HC data. This discrepancy is a potential source of bias that future researchers should be mindful of and account for in their analyses.

The protocol was approved by the Peking University People's Hospital Ethics Committee (Documentation ID: No. 2016PHB200), and written informed consent was obtained from all participating subjects for both the sharing of their data and their participation in the study. Participant information has been anonymised to protect their privacy. Personal identifiers have been removed, and data has been aggregated to prevent the identification of individual participants.

2,238 individuals, involving 1,034 RA patients and 1,204 healthy controls (HC), were enrolled in this study. Beijing cohort 1 registered the majority of RA (70%) and HC subjects (92%) (Fig. 1A). With a minority of subjects missing age information, data from 2,029 individuals showed that RA patients mainly consisted of female subjects aged between 40 to 69 (Fig. 1B). Of note, the female healthy controls enrolled in this cohort are young to middle-aged adults. We employed PERMANOVA to assess the relationships among age, gender, disease, and gut microbial composition. Our analysis revealed significant associations between these factors and microbial composition across various taxonomic levels. Importantly, even after adjusting for age and gender, we found that the disease remains significantly associated with the gut microbiota (Table 1).

Along with age, up to 91 demographic and clinical parameters were documented in this cohort (Fig. 1B–E). The metadata consists of seven demographic pieces of information (8%, such as gender, age, height, etc.), 22 clinical parameters (24%, including CRP, ESR, DAS28, etc.), 37 extra-articular manifestations involving six different systems or organs (41%, such as Oculopathy, thyroid disease, etc.), ten strategies of medication (11%, such as MTX, LEF, etc.), ten dietary habits (11%, involving milk, soybean, etc.) and five environmental factors (5%, including chemical reagent and radiation) (Fig. 1C–E). In the RA population, 85.6% of patients were positive for anti-CCP antibodies, and 70.3% were positive for rheumatoid factor.

In our study, 27% of patients were current smokers, while 4.5% had a history of smoking. Upon analyzing the association between smoking status and gut microbiota, we found no significant correlation between smoking and the bacterial Shannon index ($r = -0.010$, $P = 0.770$), observed OTUs ($r = -0.024$, $P = 0.505$), or microbial beta diversity ($R^2 = 0.016$, $P = 0.875$).

The datasets used in the current study are distinct and do not include any samples previously published.

Sample collection and DNA extraction. All fecal samples were collected using the same procedure in all centers. Briefly, fresh stool samples were frozen at -80°C within 24 h of receipt. To avoid batch effects, all samples were transferred to a center (PKUPH) for the following process. Genomic bacterial DNA was extracted using the QIAamp DNA Stool Mini Kit (QIAGEN, Germany) according to the manufacturer's recommendations. A stool mechanical disruption step with a bead-beater was performed per a previously described protocol⁷. DNA concentrations were determined using the Qubit[®] 3.0 fluorescent quantification kit (Thermo Fisher Scientific, Waltham, MA, USA). Extracted DNA was stored at -80°C before sequencing.

16S rRNA gene amplification and sequencing. The V3-V4 hypervariable regions of the 16S rRNA gene were amplified. PCR reactions were performed using unique fusion primers designed based on the universal primer set, 338 F (5'-GTACTCCTACGGGAGGCAGCA-3') and 806 R (5'-GTGGACTACHVGGGTWTCTAAT-3'), incorporating the Illumina adapters and a sample barcode sequence⁸. Triplicate PCR reactions were performed for each sample and visualized on 2% agarose gels (Thermo Fisher Scientific, Waltham, MA, USA). PCR amplicons were purified using the Agencourt AMPure XP kit (Beckman Coulter, Inc, Brea, CA, USA) and quantified by Qubit[®] 3.0 fluorescent quantification kit (Thermo Fisher Scientific, Waltham, MA, USA), then pooled at equimolar amounts as specified in the Illumina TruSeq Sample Preparation procedure. The amplicon library was constructed following the manufacturer's instructions (Illumina, San Diego, CA, USA) and quantified using the KAPA Library Quantification Kit KK4824 (KAPA Biosystems, Woburn, MA, USA). The completed library was



Fig. 1 Overview of Cohort Composition and Data Processing in Gut Microbiota Study of RA patients (from APRAC) and HCs. (A) Pie charts represent the distribution of RA patients and HCs recruited from six centers. Each segment reflects the number of participants and their percentage from each location. (B) Bar graphs depict the cohort's age distribution, categorized by gender and group (RA or HC). The x-axis indicates the number of individuals, while the y-axis lists the age ranges. (C) Pie chart displaying the age at disease onset for RA patients, with segments showing the number of patients and their corresponding age ranges. (D) Pie chart detailing the types and proportions of parameters documented by participants, categorized into demographic information, environmental factors, dietary habits, medication, clinical parameters, and extra-articular manifestations. (E) Pie chart illustrating the different extra-articular manifestations documented in the study, each segment indicates a system or organ involved and its percentage. (F) Flowchart summarizing the steps and tools used in data processing, from sampling and sequencing through bioinformatic analysis. Each icon represents a stage in the workflow, with a brief description underneath.

Taxonomic level	Factor	Sample size (HC, RA)	R2 (effect size)	P-value (Permutaion = 999)
Phylum	Disease	2038 (1086, 952)	0.00246	0.007
	Age	2038 (1086, 952)	0.07506	0.001
	Gender	2038 (1086, 952)	0.00096	0.119
	Disease (adj. age)	2038 (1086, 952)	0.07698	0.001
Class	Disease	2038 (1086, 952)	0.00407	0.001
	Age	2038 (1086, 952)	0.06352	0.001
	Gender	2038 (1086, 952)	0.00168	0.015
	Disease (adj. age&gender)	2038 (1086, 952)	0.06677	0.001
Order	Disease	2038 (1086, 952)	0.00418	0.001
	Age	2038 (1086, 952)	0.06333	0.001
	Gender	2038 (1086, 952)	0.0016	0.015
	Disease (adj. age&gender)	2038 (1086, 952)	0.06662	0.001
Family	Disease	2038 (1086, 952)	0.00401	0.001
	Age	2038 (1086, 952)	0.05714	0.001
	Gender	2038 (1086, 952)	0.00192	0.003
	Disease (adj. age&gender)	2038 (1086, 952)	0.0611	0.001
Genus	Disease	2038 (1086, 952)	0.00829	0.001
	Age	2038 (1086, 952)	0.0498	0.001
	Gender	2038 (1086, 952)	0.0018	0.001
	Disease (adj. age&gender)	2038 (1086, 952)	0.05702	0.001
ASV	Disease	2038 (1086, 952)	0.00674	0.001
	Age	2038 (1086, 952)	0.04491	0.001
	Gender	2038 (1086, 952)	0.00108	0.001
	Disease (adj. age&gender)	2038 (1086, 952)	0.05028	0.001

Table 1. The effect size of metadata variables on gut microbial composition.

sequenced on an Illumina MiSeq (Illumina, San Diego, CA, USA) platform using a dual-index sequencing strategy according to the Illumina recommended protocol⁹.

The workflow for data processing is presented in Fig. 1F. In brief, 16S rRNA gene reads were compiled and processed using the Quantitative Insights Into Microbial Ecology (QIIME Version 2, <http://www.qiime.org>) pipeline¹⁰. Raw sequences were processed to concatenate reads into tags according to their overlapping relationship; then, reads belonging to each sample were separated with barcodes, and low-quality reads (Q-score < 20) were removed. After de-noising processing by DADA2¹¹, the clean tags were clustered *de novo* into amplicon sequence variants (ASVs) with a 100% similarity of merged reads. Instead of “feature”, the term “ASV” was used throughout the manuscript. ASVs were assigned to the different taxa by matching them to the Greengenes database (Release 13.8, <https://greengenes.secondgenome.com>)¹² and chimeras were removed using the q2-feature-classifier plugin. The resulting rarefied ASV tables and their corresponding taxonomy profiles were used as input for downstream analyzes. All bioinformatic and statistical analyses were conducted in R software, version 4.1.3 (R Foundation for Statistical Computing, Vienna, Austria) unless otherwise stated.

Data Records

Raw fastq data. To share the original sequencing data for other researchers to verify, reprocess, and re-analyze according to their analytical needs and customized parameters, we deposited our raw 16S rRNA V3-V4 region amplicon data (FASTQ format) on the Genome Sequence Archive (GSA) platform¹³. This dataset is entitled “Fecal 16S rRNA genomics in patients with rheumatoid arthritis” and is available for download under GSA ID CRA003232 (<https://ngdc.cncb.ac.cn/gsa/browse/CRA003232>)¹⁴. The dataset contains 2,238 samples with 4,476 FASTQ files (two files per sample, representing forward and reverse reads). The pipeline used for the bioinformatic analysis can be downloaded through Figshare¹⁵. Other original data, such as metadata and taxonomic abundance datasets, are accessible through the R-data files tagged with the “.rds” extension.

Sample metadata. The demographic and clinical information of the study population is documented in the R-data file named “1.sample.info.original.rds”. As described previously, the metadata consists of seven demographic pieces of information, 22 clinical parameters, 37 extra-articular manifestations (EAMs) involving six different systems or organs, ten strategies of medication, ten dietary habits, and five environmental factors (Fig. 1C–E).

Taxonomic abundance files and quality report. To facilitate reanalysis for other researchers, we have shared the profiled ASV table, entitled “1.16S.ASV.profile.original.rds”, on the Figshare platform under the folder named “data”(https://figshare.com/articles/dataset/Data_for_publication_in_Scientific_Data/27603876/2)¹⁵. Additionally, we have provided the taxonomy data (“1.taxonomy.info.rds”) and the annotated bacterial ASV matrix with an abundance over 0.1% (“2-1a.tax_6Genus0.001_HC” and “2-1a.tax_6Genus0.001_RA”). The files

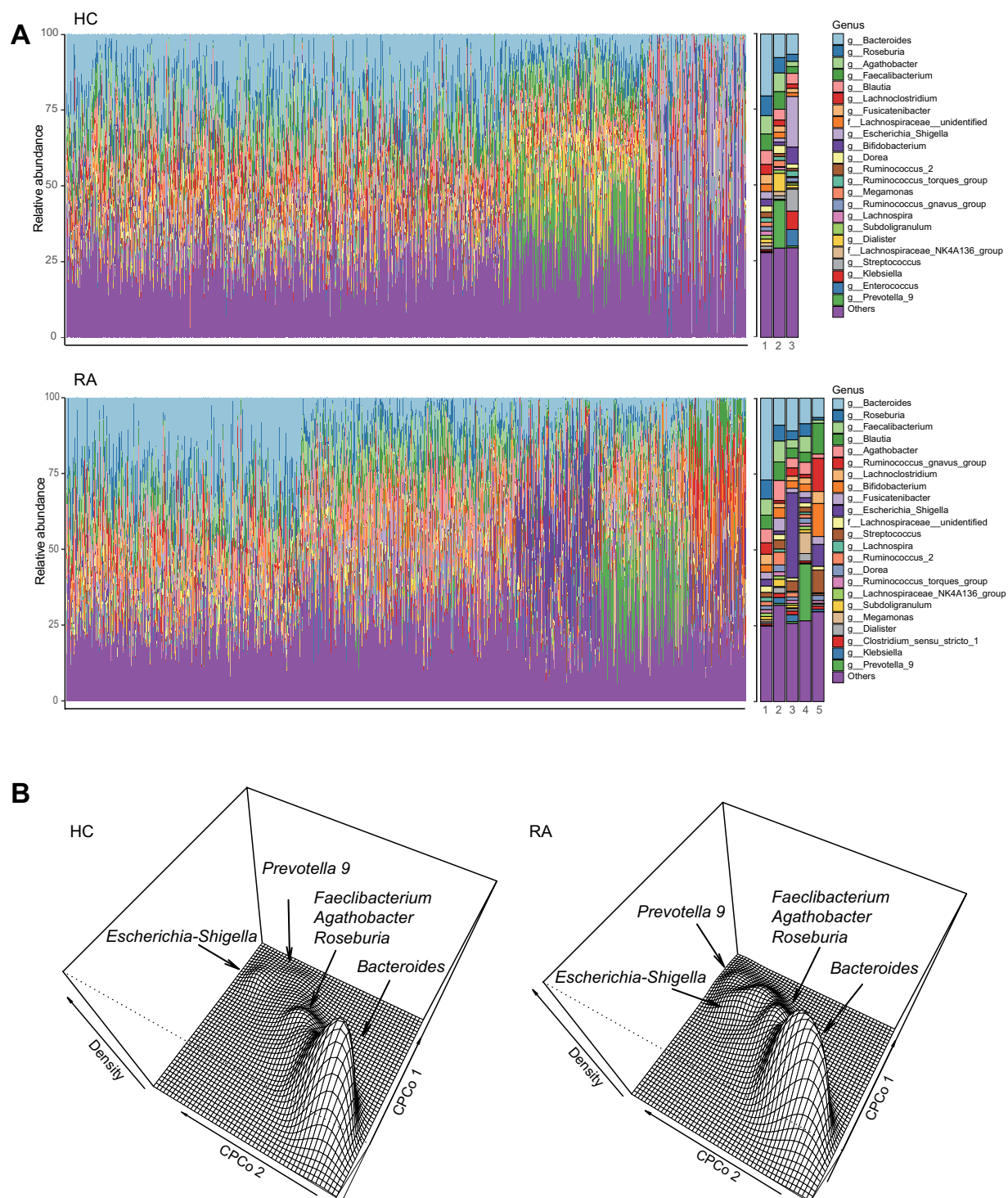


Fig. 2 Community Structure of Gut Microbiota in RA and HC Groups. **(A)** The graphs display the relative abundance of the most prevalent genera in the gut microbiota of RA patients and HCs. Each bar represents an individual's microbiota profile, with colors indicating different genera. Genera that do not fall into the top prevalent categories are grouped under 'Others.' The color-coded sidebar to the right of the plots indicates three major clusters in HC and five in RA, illustrating variations within the community composition. **(B)** The 3D PCoA plots illustrate the two groups' overall gut microbiota composition and beta diversity based on genus-level profiles, using Bray-Curtis dissimilarity. The shape of the surfaces represents the density of points in the CPCoA space, with key genera labeled to show their relative influence on the separation between HCs and RA patients.

named “stackplot” represent taxonomic data for the top 23 genera with highest abundance in both HC and RA groups and individuals. The files entitled “cuttree” were the clustering lists for two groups.

Technical Validation

Cohort control. Though individuals were enrolled from six different centers, operators in all centers followed parallel inclusion and exclusion criteria, strict documentation of epidemiological information, and fecal sample collection processes.

Sample and data. All fecal samples were collected and sent to the same center (PKUPH) for further processing. Bacterial DNA extraction was performed using the QIAamp DNA Stool Mini Kit (QIAGEN, Germany) according to the manufacturer’s recommendations, and the Thermo NanoDrop One instrument (Thermo Fisher Scientific, MA, USA) was used for controlling the quality and concentration of extracted DNA. Following amplification of 16S rRNA V3-V4 regions, PCR product quality was validated through 1% agarose gel electrophoresis and Qubit Fluorometer testing (Thermo Fisher Scientific, MA, USA). Raw sequencing data quality was checked using FastQC v0.11.9¹⁶ and MultiQC v0.4¹⁷ software, showing an average sequence count of 64,903 per sample, along with an average Q-score of 38, an average read length of 446 bp, and 52% GC content per sequence (https://figshare.com/articles/dataset/Data_for_publication_in_Scientific_Data/27603876/215, Fig. S2). QIIME (version 2) pipeline, q2-feature-classifier plugin, was used for assignment with referring to Greengenes database (Release 13.8). ASVs with a relative abundance of over 0.1% were kept. As a result, 4,567 ASVs involving 10 phyla, 16 classes, 26 orders, 57 families, and 204 genera were identified and included in further bioinformatic analysis.

Based on profiling, the “q2-vsearch” plugin in QIIME2 pipeline was used for *de novo* clustering, with a setting of 0.99 for the parameter of “p-perc-identity”. Consequently, 3 enterotypes of HC subjects and 5 RA patients were identified (Fig. 2A). In both HC and RA subjects, the genera *Bacteroides*, *Roseburia*, *Escherichia-Shigella*, and *Prevotella*_9 dominated the gut bacterial communities. Of note, though enterotype 1 in both HC and RA individuals was mainly constituted by *Bacteroides*, the abundance of *Bacteroides* was enhanced by approximately 6.5% in the RA group compared to the HC group, with an abundance of 27.0% versus 20.5%. Additionally, there was also a ~2.3% increase in the *Ruminococcus gnavus* group (3.76% versus 1.43% of abundance) in the RA group (34.6% versus 65.1% of individuals). In addition, the HC clusters were characterized by a lower abundance of specific enterotype drivers, such as *Prevotella*_9-enriched cluster 2 (15.8% of abundance and 20.1% of individuals) and *Escherichia-Shigella*-enriched 3 (16.7% of abundance and 14.6% of individuals). On the contrary, RA clusters were characterized by higher specific drivers, including *Escherichia-Shigella*-enriched cluster 3 (28.1% of abundance and 12.4% of individuals), *Prevotella*_9-enriched cluster 4 (18.8% of abundance and 12.8% of individuals), and *Bifidobacterium*-enriched cluster 5 (10.9% of abundance and 8.4% of individuals) (Fig. 2A). To infer the distribution of microbial data based on limited samples, we also performed a Kernel density estimation (KDE) with the “MASS” package¹⁸. The results also confirmed the data acquired from clustering analysis (Fig. 2B).

Data sharing. The raw data concerning the 16S rRNA gene sequencing reported in this manuscript have been deposited in the Genome Sequence Archive at the National Genomics Data Center, Beijing Institute of Genomics (China National Center for Bioinformation), Chinese Academy of Sciences, under accession number CRA003232 that are publicly accessible at <https://bigd.big.ac.cn/gsa>.

Code availability

The method descriptions list all software versions used. The codes for data analysis are available in GitHub platform at <https://github.com/JerryHnuPKUPH/1000RA>. The data analysis pipeline for this study is available on the Figshare platform at https://figshare.com/articles/dataset/Data_for_publication_in_Scientific_Data/27603876/215.

Received: 13 May 2024; Accepted: 3 January 2025;

Published online: 07 February 2025

References

1. Zaiss, M. M., Joyce Wu, H. J., Mauro, D., Schett, G. & Ciccica, F. The gut-joint axis in rheumatoid arthritis. *Nat Rev Rheumatol* **17**, 224–237 (2021).
2. Holers, V. M. *et al.* Mechanism-driven strategies for prevention of rheumatoid arthritis. *Rheumatol Autoimmun* **2**, 109–119 (2022).
3. Ahsan, H. Origins and history of autoimmunity—A brief review. *Rheumatol Autoimmun* **3**, 9–14 (2022).
4. Chu, C. Preventing rheumatoid arthritis: Lessons from that of type 1 diabetes. *Rheumatol Autoimmun* **3**, 67–69 (2023).
5. Ruff, W. E., Greiling, T. M. & Kriegel, M. A. Host-microbiota interactions in immune-mediated diseases. *Nat Rev Microbiol* **18**, 521–538 (2020).
6. He, J. *et al.* Intestinal butyrate-metabolizing species contribute to autoantibody production and bone erosion in rheumatoid arthritis. *Sci Adv* **8**, eabm1511 (2022).
7. Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027–1031 (2006).
8. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *P Natl Acad Sci USA* **108**, 4516–4522 (2011).
9. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Appl Environ Microb* **79**, 5112–5120 (2013).
10. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**, 852–857 (2019).
11. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**, 581 (2016).
12. DeSantis, T. Z. *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microb* **72**, 5069–5072 (2006).

13. Chen, T. *et al.* The Genome Sequence Archive Family: Toward Explosive Data Growth and Diverse Data Types. *Genomics, Proteomics & Bioinformatics* **19**, 578–583 (2021).
14. Li J *et al.* A Comprehensive Dataset on Microbiome Dynamics in Rheumatoid Arthritis from a Large-Scale Cohort Study. *Genome Sequence Archive, Dataset*. <https://ngdc.cncb.ac.cn/gsa/search?searchTerm=CRA003232> (2024).
15. Li, J. *et al.* A comprehensive dataset on microbiome dynamics in rheumatoid arthritis from a large-scale cohort study. *Figshare, Dataset* <https://doi.org/10.6084/m9.figshare.27603876.v2> (2024).
16. Wingett, S.W. & Andrews, S. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research* **7** (2018).
17. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
18. Falony, G. *et al.* Population-level analysis of gut microbiome variation. *Science* **352**, 560–564 (2016).

Acknowledgements

This study was supported by Shenzhen Medical Research Fund (C2404002), Natural Science Foundation of China (92374202, 82271835, 32470956, 82370555, 32441099, 32141004, 81901648), Sanming Project of Medicine in Shenzhen (SZSM202311030), and Capital's Funds for Health Improvement and Research (CFH2024-4-4089).

Author contributions

J.H. and Z.L. conceived the study and performed the analyses. J.H., J.L., and J.X. wrote and edited the manuscript. Y.G., Y.W., R.F., W.F., Y.L., X.Z., Y.L., S.G., L.S., Y.C., L.S., X.S., Y.X., Q.W., R.L., J.Z., Y.L.L. and J.H. collected and processed the samples. J.X., C.X., and J.L. deposited the sequencing data in the databases. C.X., J.X. and J.Q. performed the bioinformatic analyses, interpreted the results, and designed figures. All the authors have revised and approved the manuscript submission.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Z.L. or J.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025