Gut: first published as 10.1136/gutjnl-2022-328034 on 1 September 2022. Downloaded from http://gut.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

Ageing trajectory of the gut microbiota is associated with metabolic diseases in a chronological age-dependent manner

We read with interest the recent article by Ng *et al*. (*Gut*, 2022; 71:910–8), who reported inhibition of the gut microbiota trajectory in patients with autism spectrum disorder.[1] Similarly, the human gut microbiota ages in adults.[2 3] Transplantation of gut microbes from elderly hosts, compared with their younger counterparts, deteriorates recipients' age-related metabolic alternations.[4 5] However, the gut microbiota in humans could differ in its pace of ageing, namely, accelerated or delayed microbiota ageing, even among those with similar chronological ages (figure 1A); this process is analogous to biological ageing.[6] Therefore, we wondered whether the gut microbiota ageing trajectory could be used as a biomarker for metabolic diseases in adults. We analysed the gut microbiota compositions of 6376 participants of a population-level survey[7] with Quantitative Insights Into Microbial Ecology (QIIME),[8] among whom the median chronological age of the healthy participants is 43 (IQR, 31–45, online supplemental table S1). A random forest algorithm was applied to the data of 1083 healthy individuals to model gut microbiota ageing in relation to chronological age. The model was then applied to subjects with metabolic

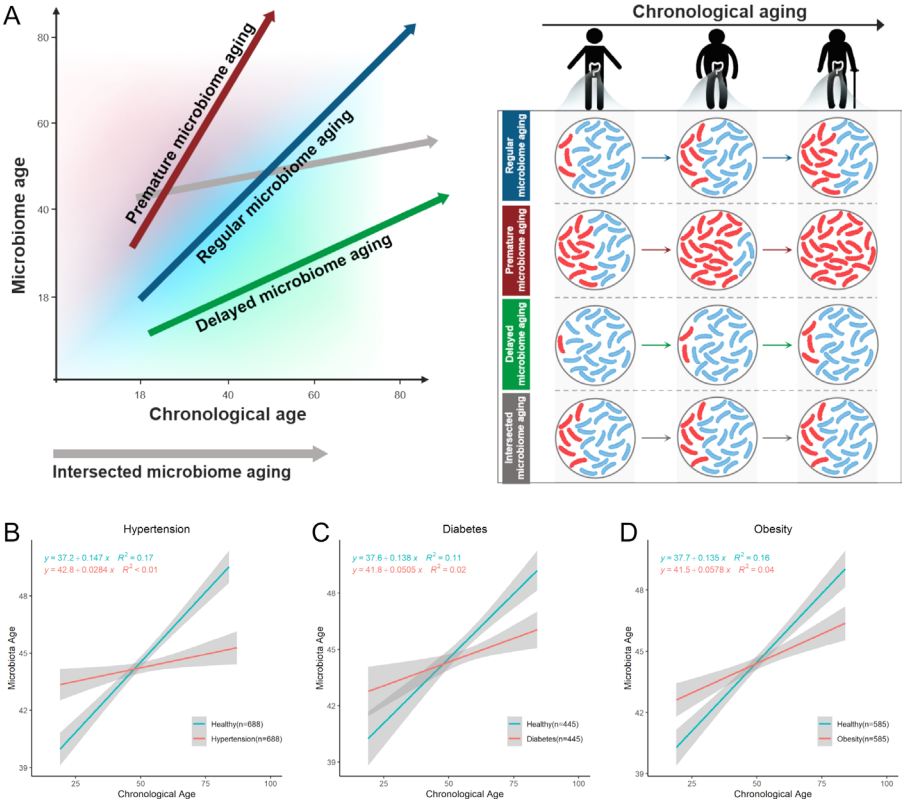

Figure 1 Gut microbiome ageing patterns in participants with metabolic diseases and their healthy counterparts. (A) Three different patterns of ageing trajectories of the gut microbiota. (B–D) Gut microbiota ageing trajectory in patients with hypertension, diabetes and obesity versus healthy individuals. The random forest model was trained with the microbiota features of healthy individuals and then applied to predict the microbiota age of participants with hypertension, diabetes and obesity in relation to their chronological their age separately.

diseases, matching healthy individuals by chronological age. See the online supplemental file 1 for additional information.

At the modelling phase, the microbiota age of the 1083 healthy individuals showed a significant and positive correlation with chronological age (online supplemental figure S1). The microbiota age in the healthy individuals was not significantly associated with gender or residing sites (developed vs less developed), but is significantly associated with grain, fruits and rice wine consumption (online supplemental table S2 and online supplemental figure S2). When the model was applied to individuals with metabolic diseases, the ageing trajectory of their microbiota was neither accelerated nor delayed but intersected with that of healthy individuals at nearly 50 years old (figure 1B–D). In individuals with metabolic disease, the microbiota age was older than that in healthy individuals at younger chronological ages, comparable with that in healthy individuals at middle chronological ages, and younger than that in healthy individuals at older chronological ages (online supplemental figure S3). Metabolic diseases (including hypertension, diabetes and metabolic syndrome) patients who took antibiotics or medication showed significantly lower microbiota age than the treatment-naive patients (online supplemental figure S4). We removed the medicated patients and the above intersecting patterns could still be reproduced (online supplemental figure S5).

To dissect the random forest ageing model, we analysed the ageing trajectories of the top 20 model-contributing taxa (figure 2A). The heatmap clearly showed that the relative taxonomic abundances correlated positively (upper half) or negatively (lower half) with chronological ageing in the healthy individuals, but the trajectories were disordered in hypertensive individuals. Typical microbial examples include *Clostridium* and *Parabacteroides distasonis*, with obvious intersecting patterns (figure 2B,C). Similar disordered taxonomic ageing trajectories were also observed in participants with diabetes, obesity, metabolic syndrome, hypercholesterolemia and hypertriglyceridaemia (online supplemental figure S6).

In conclusion, the results of the current and Ng's study suggest that the ageing trajectory of the gut microbiota

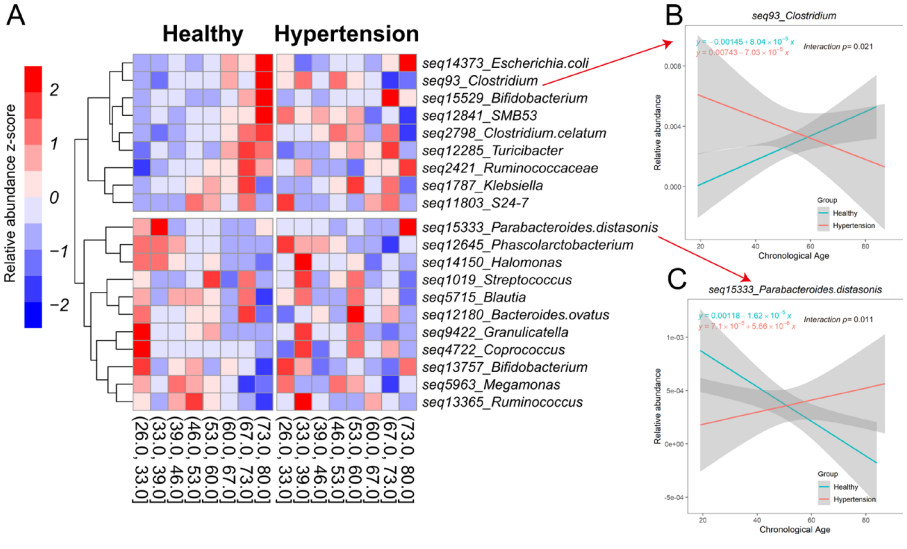

Figure 2 Significant intersecting trajectories of taxonomic relative abundances in healthy and hypertensive participants. (A) Heatmap of the top 20 important amplicon sequence variant from the random forest regression model plotted against chronological age in healthy and hypertensive individuals. Significant intersecting patterns of the relative abundances of (B) *seq93_Clostridium* and (C) *seq15333_Parabacteroides distasonis* in healthy individuals and hypertensive participants. The interaction p indicates the significance of interaction term in the multivariable linear regression model.

Gut: first published as 10.1136/gutjnl-2022-328034 on 1 September 2022. Downloaded from http://gut.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

could be a potential biomarker for both paediatric and adult chronic diseases. However, accelerated microbiota ageing should not be simply considered a risk factor for metabolic diseases in adults or vice versa, as it might be chronological-age dependent. Our finding corresponds to Wilmanski's observation that elderly individuals who carried younger microbiota signatures could have lower survival rates,[9] but analyses to understand covariates (lifestyle, diet, medication, etc), subpopulation differences and longitudinal disease risks of gut microbiota ageing are warranted. Moreover, whether a chronological age-dependent bacterial function could be observed and understood in mechanism studies is worth investigating. Interestingly, a common practice in faecal microbiota transplantation (FMT) is to choose young donors for health reasons. Analysing whether a chronological age-matched healthy donor for FMT could promote additional health benefits in the recipient is worth further investigation.

JingXiang Fu ,[1] **Wen Qiu,**[1] **Huimin Zheng,**[1] **Cancan Qi,**[1] **Shixian Hu** ,[2] **Wei Wu,**[3] **Huidi Wang,**[1] **Guangyan Wu,**[1] **Peihua Cao,**[4,5] **Zhenchao Ma,**[6] **Chao Zheng,**[6] **Wen-Jun Ma,**[7] **Hong-Wei Zhou** ,[1,8] **Yan He**[1,8]

[1]Microbiome Medicine Center, Division of Laboratory Medicine, Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong, China
[2]Institute of Precision Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China
[3]Guangdong Provincial Institute of Public Health, Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China
[4]Clinical Research Center, Zhujiang Hospital, Southern Medical University, Guangzhou, China
[5]Department of Biostatistics, School of Public Health, Southern Medical University, Guangzhou, China
[6]Huzhou Central Hospital, School of Medicine, Zhejiang University, Huzhou, Zhejiang, China
[7]Department of Public Health and Preventive Medicine, School of Medicine, Jinan University, Guangzhou, Guangdong, China
[8]State Key Laboratory of Organ Failure Research, Southern Medical University, Guangzhou, China

Correspondence to Professor Yan He, Microbiome Medicine Center, Division of Laboratory Medicine, Zhujiang Hospital, Southern Medical University, Guanghzou, China; yanhe@i.smu.edu.cn

Correction notice This article has been corrected since it published Online First. The first author's name in the citationof reference 9 has been corrected.

Acknowledgements We thank Professor Lianmin Chen from Nanjing Medical University for an insightful discussion. YH (corresponding author) argued with Professor Lianmin Chen about whether a younger gut microbiota is always a healthier one, and then initiated the present analysis, in addition to the inspiration by Ng's study. Based on the present observations, YH could be wrong for insisting that a younger microbiota is always better, but might be chronological-age dependent.

Contributors JF, YH and H-WZ designed the study and prepared the manuscript. YH, WW, W-JM and H-WZ provided the data. JF, WQ, H-MZ, HW, GW and YH analysed the data. PC, ZM and CZ provided crucial advice in analysing and interpreting the data.

Funding This study was supported by the National Key R&D Program of China (2019YFA0802300 (YH), 2017YFC1310600 (H-WZ)); National Natural Science Foundation of China (NSFC82022044 (YH), NSFC81925026 (H-WZ), 81800746 (YH)); and Natural Science Foundation of Guangdong Province China, Project No.-2018A0303130118 (WW).

Competing interests None declared.

Patient consent for publication Consent obtained directly from patient(s)

Ethics approval This study involves human participants and was approved by the ethical review committee of the Chinese Center for Disease Control and Prevention (no. 201519-A). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

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

OPEN ACCESS

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

© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2022-328034).

JF and WQ contributed equally.

Check for updates

To cite Fu J, Qiu W, Zheng H, *et al*. *Gut* 2023;**72**:1431–1433.

Received 9 June 2022
Accepted 14 August 2022
Published Online First 1 September 2022

Gut 2023;**72**:1431–1433. doi:10.1136/gutjnl-2022-328034

ORCID iDs
JingXiang Fu http://orcid.org/0000-0003-3009-5269
Shixian Hu http://orcid.org/0000-0002-1190-0325
Hong-Wei Zhou http://orcid.org/0000-0003-2472-8541

REFERENCES

1 Wan Y, Zuo T, Xu Z, *et al*. Underdevelopment of the gut microbiota and bacteria species as non-invasive markers of prediction in children with autism spectrum disorder. *Gut* 2022;71:910–8.
2 Yatsunenko T, Rey FE, Manary MJ, *et al*. Human gut microbiome viewed across age and geography. *Nature* 2012;486:222–7.
3 Zhang X, Zhong H, Li Y, *et al*. Sex- and age-related trajectories of the adult human gut microbiota shared across populations of different ethnicities. *Nat Aging* 2021;1:87–100.
4 Bárcena C, Valdés-Mas R, Mayoral P, *et al*. Healthspan and lifespan extension by fecal microbiota transplantation into progeroid mice. *Nat Med* 2019;25:1234–42.
5 Boehme M, Guzzetta KE, Bastiaanssen TFS, *et al*. Microbiota from young mice counteracts selective age-associated behavioral deficits. *Nat Aging* 2021;1:666–76.
6 Hamczyk MR, Nevado RM, Barettino A, *et al*. Biological versus chronological aging: JACC focus seminar. *J Am Coll Cardiol* 2020;75:919–30.
7 He Y, Wu W, Zheng H-M, *et al*. Regional variation limits applications of healthy gut microbiome reference ranges and disease models. *Nat Med* 2018;24:1532–5.
8 Caporaso JG, Kuczynski J, Stombaugh J, *et al*. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;7:335–6.
9 Wilmanski T, Diener C, Rappaport N, *et al*. Gut microbiome pattern reflects healthy ageing and predicts survival in humans. *Nat Metab* 2021;3:274–86.

Gut: first published as 10.1136/gutjnl-2022-328034 on 1 September 2022. Downloaded from http://gut.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

Online methods

Study participants

In our previously reported Guangdong Gut Microbiome Project[1], we elaborated on the details of the participants and survey methods. Briefly, samples that were selected randomly under a standard protocol using a probability proportional to size (PPS) sampling method were collected from 14 different districts in Guangdong Province[2]. The 14 districts were divided into developed and less developed city groups, with Guangzhou, Foshan, and Shenzhen constituting the developed city group and the rest of the cities being classified as the less developed city groups. After removing the samples obtained from those who did not complete the questionnaire and those with faecal sample sequences of less than 10,000 reads. In addition, a total of 6,376 samples were retained in this study after a preprocessing procedure was performed on the data. We further selected 1083 healthy individuals from the cohort as healthy controls. The inclusion criteria for healthy individuals were as follows: no reported illness; fasting blood glucose (FBG) levels of <6.1; BMI of <24; no antibiotic use within 1 month of stool sample donation; no clinical parameters indicating diagnoses of hypertension, hypertriglyceridaemia, hypercholesterolemia, neurosis, or chronic fatigue syndrome (CFS); and no medication for metabolic diseases intake before donating the stool sample. As the age composition differed significantly between the healthy control group and the different disease groups, we age-matched the different disease groups to the healthy control group using chronological age to avoid its effect as a strong confounding factor. The matched disease groups were as follows: hypertriglyceridaemia (n=537), hypercholesterolemia (n=571), metabolic syndrome (MetS) (n=695), diabetes (n=445), hypertension (n=688) and obesity (n=585).

Sample processing

We described the sample processing method in our previous study[1]. In brief, we extracted total bacterial DNA from stool samples. The barcoded primers (5′ to 3′) by which the 16S rRNA gene V4 region was amplified were V4F, GTGYCAGCMGCCGCGGTAA and V4R, GGACTACNVGGGTWTCTAAT. The PCR amplification conditions were as follows: initial denaturation at 94 °C for 5 min, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 52 °C for 30 s, extension at 72 °C for 45 seconds, and a final elongation at 72 °C for 5 min. Then, the PCR products were sent to Beijing Genome Institute (BGI; Beijing, China), and next-generation sequencing was performed on an Illumina HiSeq 2500 platform.

Bioinformatics and biostatistics

The method used for the preprocessing of the raw sequences was described in our previous study[1]. Briefly, using QIIME V.1.9.1, we merged the data, performed quality control assessments and analysed the data at the sequence level using the reference-free statistical denoising deblur method to prevent improperly identifying distinct operational taxonomic unit-based data originating from misclustered sequences[2, 3].

To establish a microbiome model to predict the age of the host's microbiota, we adopted the relative abundance data at the ASV level of 1083 relatively healthy people to train a random forest regression model by using the R packages ‘randomForest’ (version 4.6.12) and ‘caret’ (version 6.0.68) to tune hyperparameters and perform 10-fold cross-validation. The sample matching process was performed using the R package 'MatchIt'. It used the propensity score method (PSM) and greedy nearest neighbour matching to achieve pairwise matching of participants in different disease groups with their healthy control counterparts. Spearman's rank correlation test was used to analyse the correlation between amplicon sequence variant (ASVs) and host chronological age, and then $P \leq 0.05$ was used as the threshold to filter out the ASVs related to chronological age for modeling. The interaction effect between chronological age and different metabolic disorders was indicated by the significance of the interaction term in the multivariable linear regression model (Relative abundance of the ASV or the microbiota age ~ chronological age + disease group + chronological age*disease group). All plots in this study were drawn using the R package ‘ggplot2’ and ‘pheatmap’, and the statistical indicators in the plots were calculated using the R packages ‘ggpubr’ and ‘ggpmisc’.

Ethics approval

Ethical approvals were described previously[1]. The Guangdong gut microbiome project was approved by the ethical review committee of the Chinese Center for Disease Control and Prevention (No. 201519-A). Written consent was obtained from all participants.

Patient and public involvement

Neither patients nor the public were involved in the design, conduction, reporting, or dissemination plans of our research.

References for online methods

[1] Y. He, W. Wu, H.-M. Zheng, P. Li, D. McDonald, H.-F. Sheng, M.-X. Chen, Z.-H. Chen, G.-Y. Ji, Z.-D.-X. Zheng, P. Mujagond, X.-J. Chen, Z.-H. Rong, P. Chen, L.-Y. Lyu, X. Wang, C.-B. Wu, N. Yu, Y.-J. Xu, J. Yin, J. Raes, R. Knight, W.-J. Ma, and H.-W. Zhou, Regional variation limits applications of healthy gut microbiome reference ranges and disease models. Nature Medicine, 2018. 24(10): 1532-1535.

[2] Lao X Q , Ma W J , Sobko T , et al. Overall obesity is leveling-off while abdominal obesity continues to rise in a Chinese population experiencing rapid economic development: analysis of serial cross-sectional health survey data 2002–2010[J]. International Journal of Obesity.

[3] J.G. Caporaso, J. Kuczynski, J. Stombaugh, K. Bittinger, F.D. Bushman, E.K. Costello, N. Fierer, A.G. Pena, J.K. Goodrich, J.I. Gordon, G.A. Huttley, S.T. Kelley, D. Knights, J.E. Koenig, R.E. Ley, C.A. Lozupone, D. McDonald, B.D. Muegge, M. Pirrung, J. Reeder, J.R. Sevinsky, P.J. Tumbaugh, W.A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, and R. Knight, QIIME allows analysis of high-throughput community sequencing data. Nature Methods, 2010. 7(5): 335-336.

[4] A. Amir, D. McDonald, J.A. Navas-Molina, E. Kopylova, J.T. Morton, Z. Zech Xu, E.P. Kightley, L.R. Thompson, E.R. Hyde, A. Gonzalez, and R. Knight, Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. mSystems, 2017. 2(2).

Supplementary tables

Table S1. Characteristics of the participants included in this study.

Characteristic	**Healthy**, N = 1,083	**Diabetes**, N = 555	**Hypercholesterolemia**, N = 775	**Hypertension**, N = 2,289	**Hypertriglyceridemia**, N = 674	**MetS**, N = 1,254	**Obesity**, N = 633	**p-value**[1]
Age, Median (IQR)	43 (31 – 54)	59 (52 – 66)	57 (49 – 65)	62 (53 – 70)	54 (45 – 63)	58 (49 – 66)	53 (44 – 62)	<0.001
Gender(Female), n (%)	638 (59)	298 (54)	431 (56)	1,194 (52)	297 (44)	601 (48)	379 (60)	<0.001
Weight(kg), Median (IQR)	53 (48 – 58)	61 (54 – 68)	58 (51 – 66)	59 (52 – 67)	64 (57 – 72)	66 (59 – 74)	74 (68 – 81)	<0.001
BMI, Median (IQR)	21.2 (19.6 – 22.5)	24.4 (22.3 – 27.1)	23.4 (21.3 – 25.9)	24.1 (21.8 – 26.5)	25.1 (22.9 – 27.6)	26.1 (24.1 – 28.4)	29.4 (28.6 – 31.2)	<0.001
Waist(cm), Median (IQR)	73 (68 – 78)	85 (79 – 92)	81 (75 – 88)	84 (77 – 90)	86 (80 – 92)	90 (85 – 95)	95 (90 – 100)	<0.001
FBG(mmol/L), Median (IQR)	5.04 (4.68 – 5.34)	8.06 (7.19 – 10.32)	5.47 (5.05 – 6.06)	5.52 (5.08 – 6.14)	5.74 (5.28 – 6.50)	6.11 (5.35 – 6.98)	5.52 (5.07 – 6.28)	<0.001
TCHO(mmol/L), Median (IQR)	5.00 (4.52 – 5.46)	5.39 (4.92 – 5.93)	6.56 (6.35 – 6.86)	5.36 (4.85 – 5.90)	5.40 (4.89 – 5.99)	5.32 (4.80 – 5.89)	5.28 (4.80 – 5.82)	<0.001
TG(mmol/L), Median (IQR)	0.81 (0.64 – 1.10)	1.53 (1.09 – 2.32)	1.30 (0.91 – 1.96)	1.22 (0.85 – 1.80)	3.15 (2.62 – 4.10)	2.06 (1.53 – 2.94)	1.48 (1.05 – 2.23)	<0.001
UA(μmol/L), Median (IQR)	304 (258 – 359)	322 (269 – 401)	329 (273 – 398)	344 (284 – 412)	390 (312 – 459)	368 (309 – 442)	365 (311 – 438)	<0.001
HDL(mmol/L), Median (IQR)	1.31 (1.10 – 1.56)	1.07 (0.89 – 1.34)	1.33 (1.13 – 1.59)	1.21 (0.99 – 1.48)	0.89 (0.73 – 1.08)	0.94 (0.81 – 1.08)	1.07 (0.87 – 1.29)	<0.001
LDL(mmol/L), Median (IQR)	2.88 (2.40 – 3.37)	3.39 (2.75 – 4.10)	4.31 (3.65 – 4.95)	3.41 (2.80 – 4.02)	3.53 (2.87 – 4.23)	3.43 (2.84 – 4.10)	3.49 (2.81 – 4.08)	<0.001
HbA1c(%), Median (IQR)	4.70 (4.30 – 5.10)	6.10 (5.20 – 7.60)	5.10 (4.60 – 5.50)	5.00 (4.60 – 5.50)	5.00 (4.60 – 5.60)	5.20 (4.70 – 5.80)	5.10 (4.60 – 5.50)	<0.001
Hb(g/L), Median (IQR)	142 (129 – 153)	143 (129 – 156)	144 (131 – 157)	143 (130 – 156)	148 (135 – 162)	147 (133 – 159)	145 (134 – 158)	<0.001
ALT(U/L), Median (IQR)	13 (10 – 18)	17 (12 – 26)	17 (12 – 23)	16 (12 – 23)	20 (14 – 30)	19 (13 – 27)	19 (14 – 29)	<0.001
BUN(mmol/L), Median (IQR)	4.82 (4.04 – 5.68)	5.36 (4.34 – 6.38)	5.42 (4.60 – 6.29)	5.39 (4.52 – 6.38)	5.08 (4.36 – 6.05)	5.13 (4.34 – 6.16)	5.03 (4.20 – 6.03)	<0.001
Antibotics use within 1 month, n (%)	0 (0)	42 (7.9)	48 (6.4)	119 (5.3)	52 (7.9)	79 (6.5)	38 (6.1)	<0.001
Medication use for metabolic diseases, n (%)	0 (0)	184 (34)	126 (17)	542 (24)	141 (22)	328 (27)	131 (21)	<0.001

[1] Kruskal-Wallis rank sum test for continuous variables; Pearson's Chi-squared test for categorical variables

Table S2. The associations of potential covariates with microbiota age in the 1083 healthy participants were measured by the multivariable linear regression model.

Covariate	Forest Plot	Beta[1]	SE[2]	95% CI[2]	p-value
Chronological Age		0.37***	0.030	0.31, 0.43	<0.001
Sleep time per day		0.02	0.028	-0.04, 0.07	0.543
BMI		-0.01	0.028	-0.07, 0.04	0.636
Grains		0.06*	0.029	0.00, 0.11	0.050
Fruits		-0.11***	0.030	-0.16, -0.05	<0.001
Fruit drinks		-0.03	0.031	-0.09, 0.03	0.410
Vegetables		0.01	0.030	-0.05, 0.06	0.839
Livestock meat		-0.01	0.030	-0.07, 0.05	0.654
Carbonated beverage		0.04	0.031	-0.02, 0.10	0.233
High alcohol liquor		0.03	0.028	-0.02, 0.09	0.237
Low alcohol liquor		0.00	0.032	-0.07, 0.06	0.922
Beer		-0.03	0.031	-0.09, 0.03	0.337
Yellow rice wine		-0.03	0.028	-0.09, 0.02	0.251
Rice wine		0.07*	0.032	0.01, 0.13	0.028
Wine		-0.05	0.031	-0.11, 0.01	0.126

[1] Beta is the standardized regression coefficient in the linear regression model.
*p<0.05; **p<0.01; ***p<0.001

[2] SE = Standard Error, CI = Confidence Interval

Supplementary figures

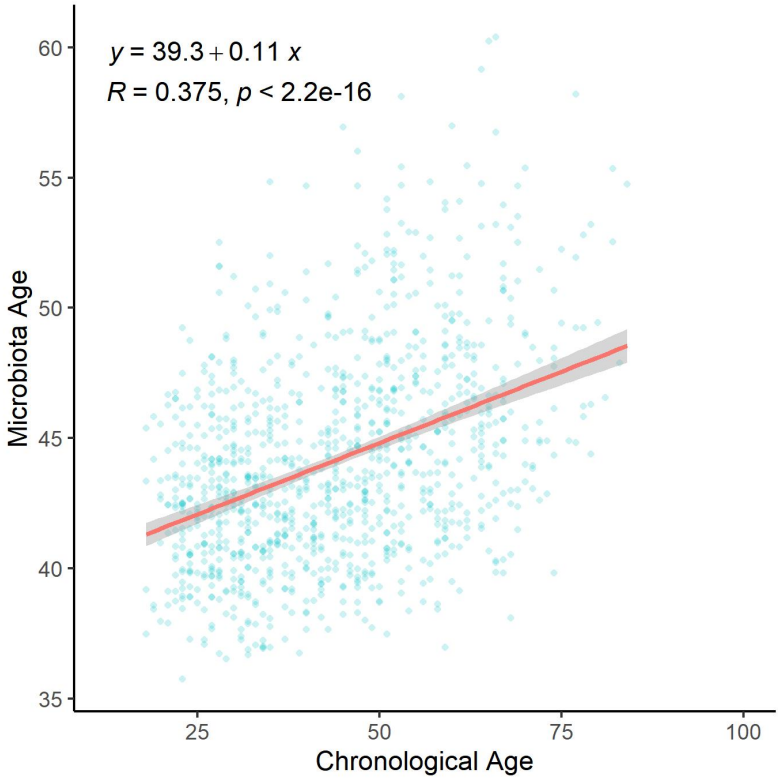

Figure S1. The scatter plot showed a significant and positive correlation between the microbiota age of 1083 healthy individuals predicted by the random forest model and their chronological age. R, Spearman correlation coefficient.

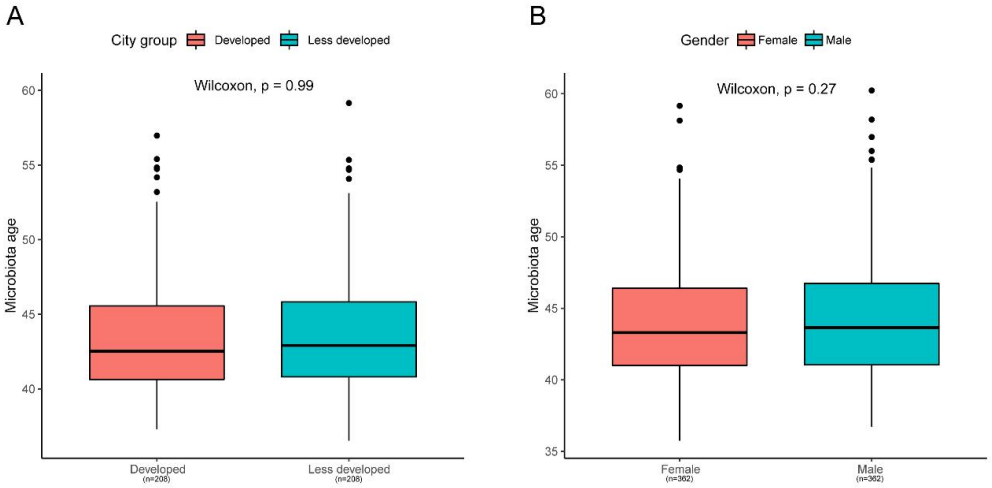

Figure S2. Comparing gut microbiota age between geographic areas (A) and gender (B) in the healthy individuals. Chronological ages were matched between different groups using the propensity score method.

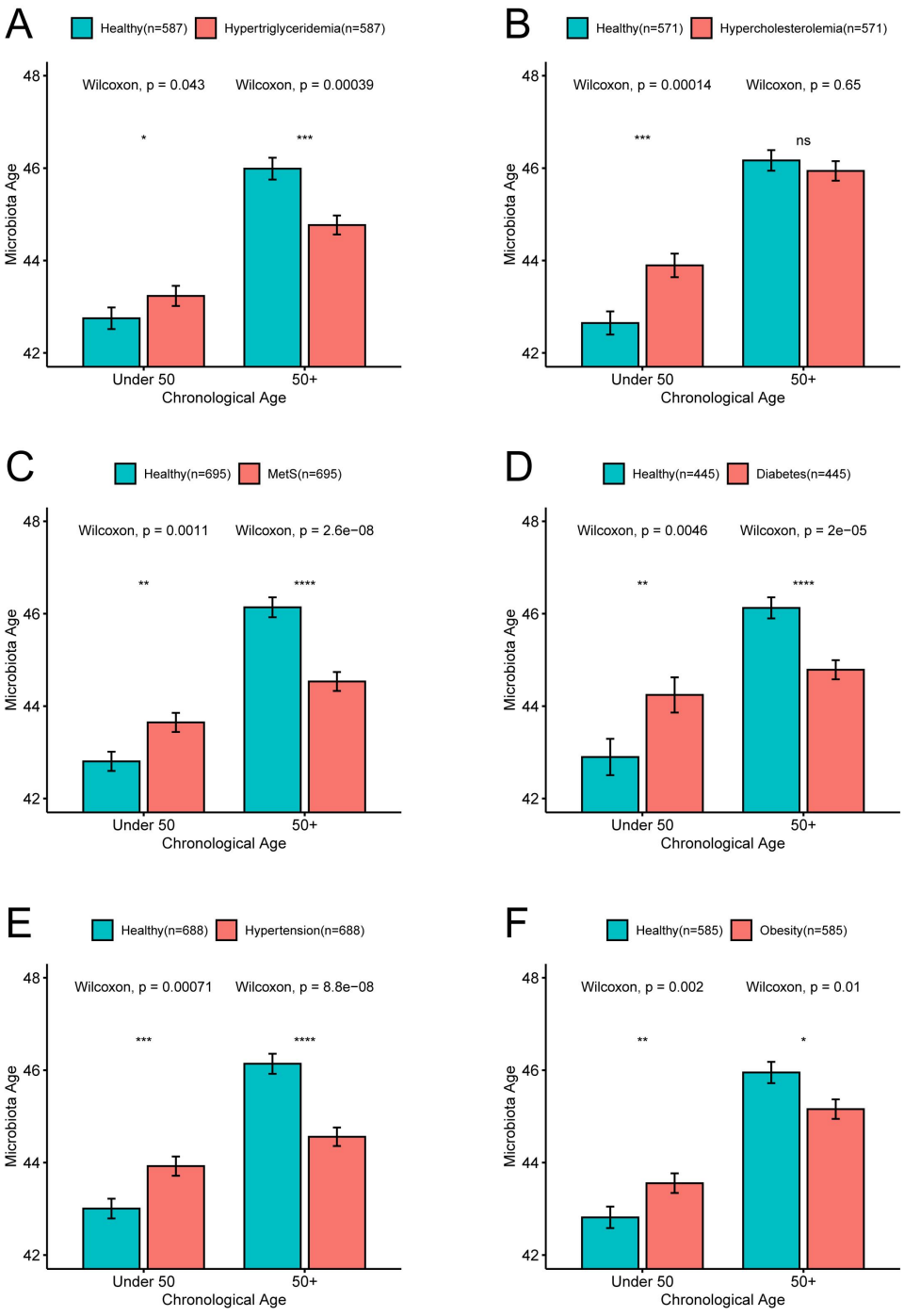

Figure S3. Comparing microbiota age between people with different metabolic disorders and healthy individuals. Figures S2(A) to S2(F) illustrate the comparisons between groups with hypertriglyceridaemia, hypercholesterolemia, metabolic syndrome (MetS), diabetes, hypertension, obesity and healthy individuals, respectively. The "under50" and "50+" labels indicate that the chronological age of the participant in their corresponding group is less than 50 years and greater than or

equal to 50 years, respectively. Significantly different groups are indicated with **** for $P \leqslant 0.0001$, *** for $P \leqslant 0.001$, ** for $P \leqslant 0.01$, * for $P \leqslant 0.05$ or ns for $P > 0.05$ (two-sided Wilcoxon test).

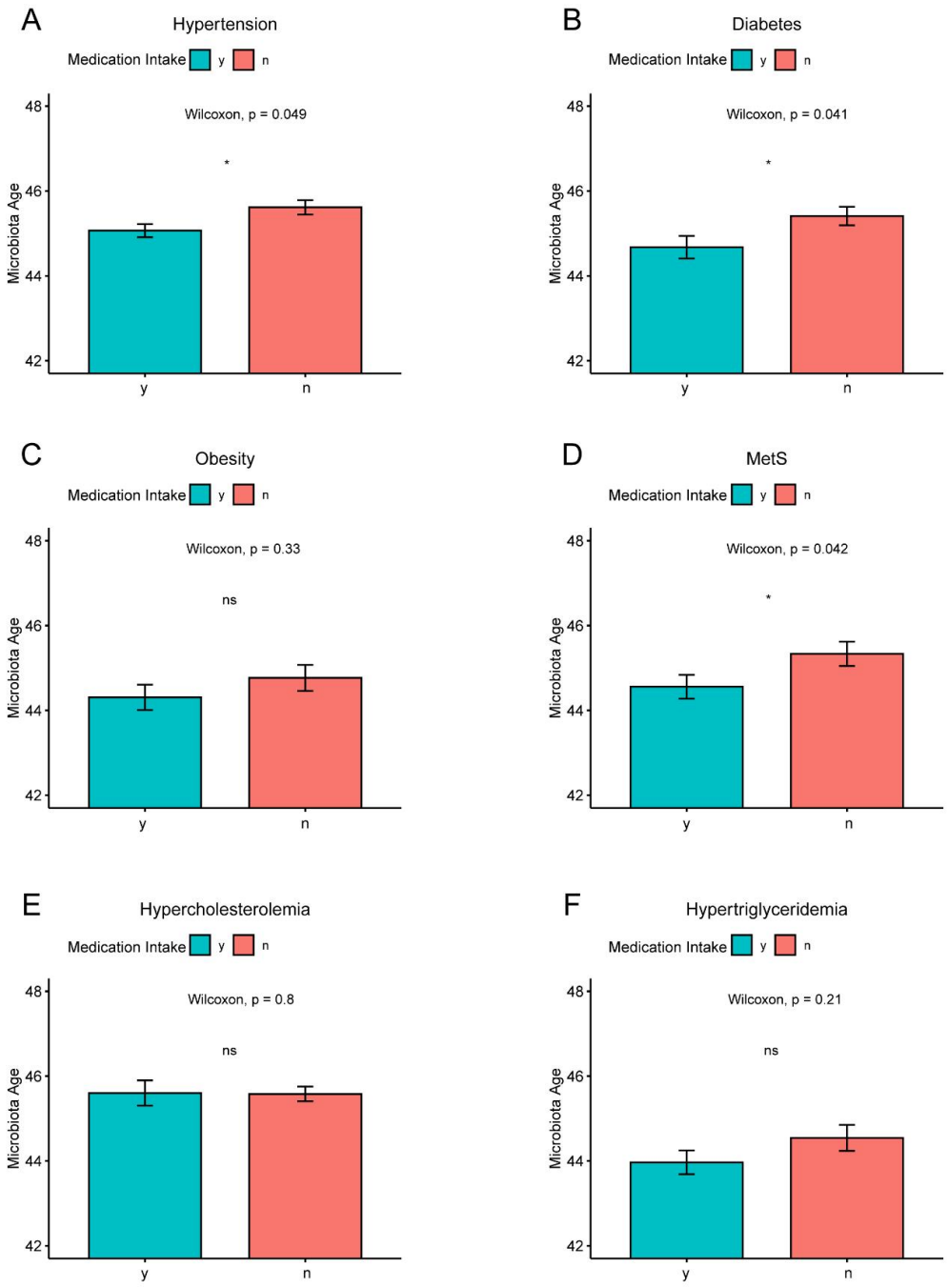

Figure S4. Comparison of gut microbiota age between metabolic diseases patients who took medication and those who are unmedicated. Figures S3(A) to S3(F) show the differences in microbiota age of individuals with hypertension, diabetes, obesity, metabolic syndrome (MetS), hypercholesterolemia and hypertriglyceridemia respectively after matching for chronological age between the two groups. Significantly different groups are indicated with **** for P⩽ 0.0001, *** for P ⩽ 0.001, ** for P ⩽ 0.01, * for P ⩽ 0.05 or ns for P > 0.05 (two-sided Wilcoxon test).

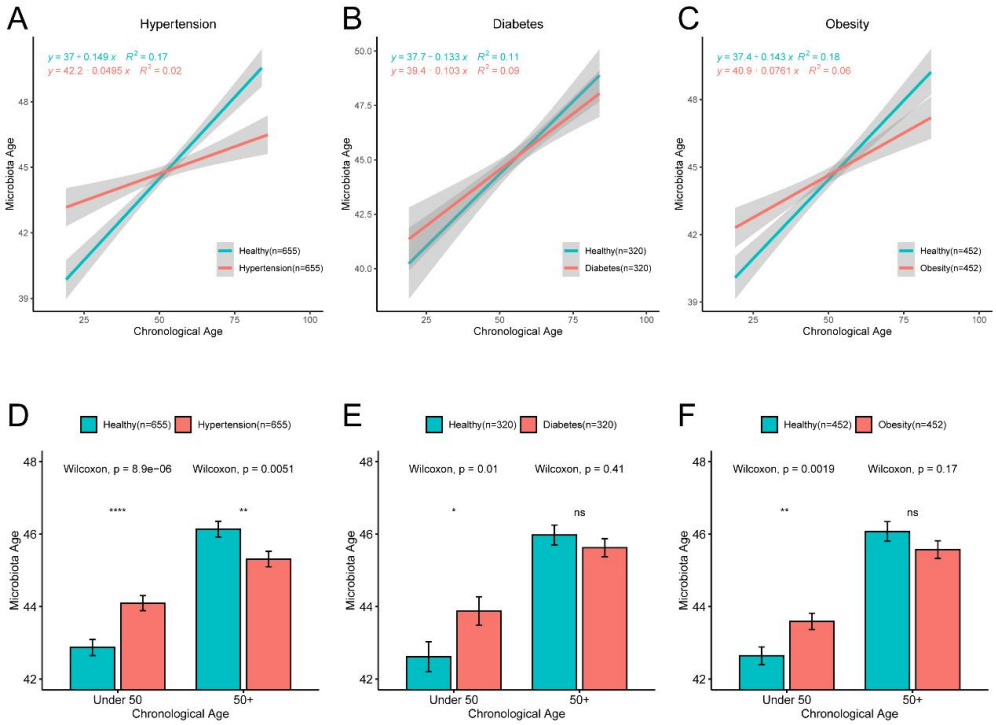

Figure S5. Reproducing the intersecting pattern between gut microbiota age and metabolic diseases after excluding medicated patients. In order to avoid the effect of antibiotic use and intake of medication on our results, we excluded subjects who used antibiotics within one month and used medication for metabolic diseases before sample collection in different disease groups. Meanwhile, patients' chronological age was matched with healthy individuals. Figures S3(A) to S3(C) illustrate the microbiota ageing trajectories of chronological age-matched participants in hypertension, diabetes and obesity respectively. Significantly different groups are indicated with **** for P⩽ 0.0001, *** for P ⩽ 0.001, ** for P ⩽ 0.01, * for P ⩽ 0.05 or ns for P > 0.05 (two-sided Wilcoxon test).

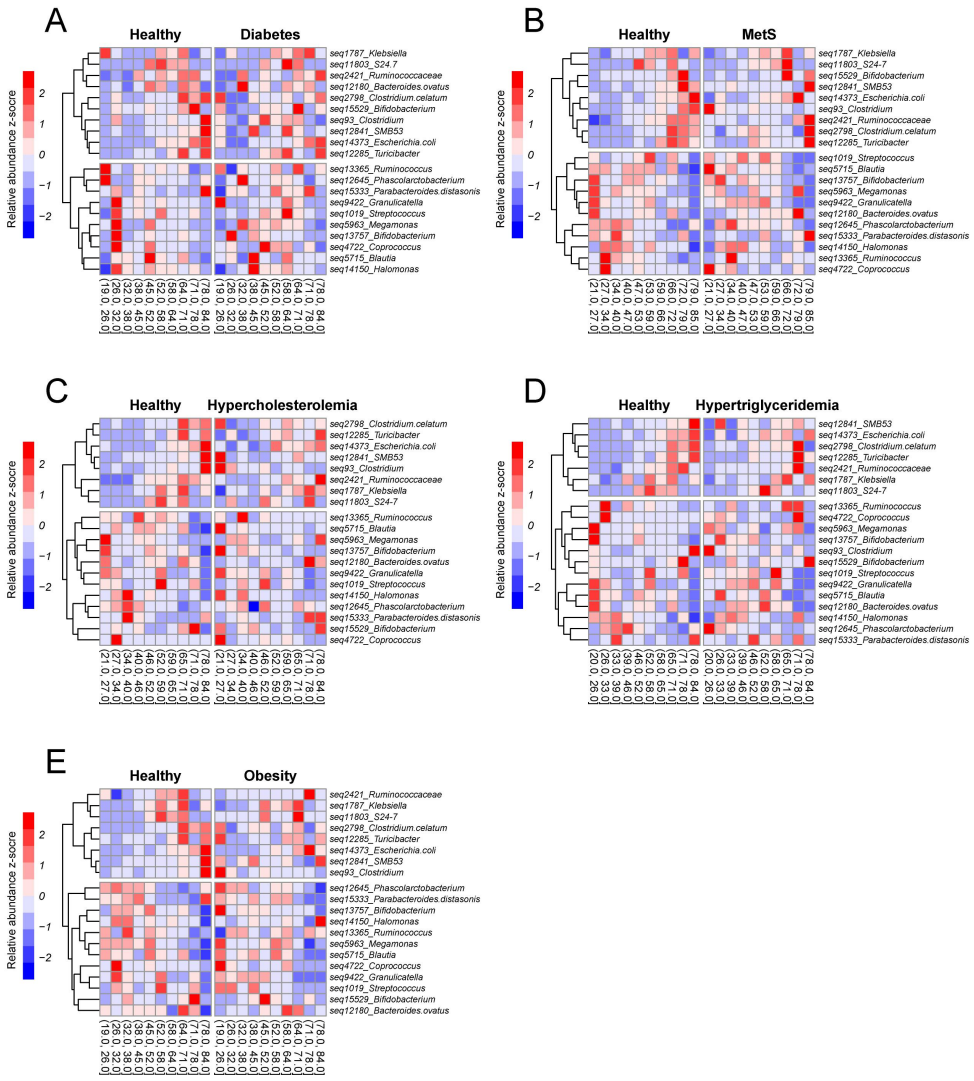

Figure S6. The aging trajectories of gut microbes in participants with different metabolic diseases and healthy individuals. The healthy and diseased groups were matched by chronological ages.

Online methods

Study participants

In our previously reported Guangdong Gut Microbiome Project[1], we elaborated on the details of the participants and survey methods. Briefly, samples that were selected randomly under a standard protocol using a probability proportional to size (PPS) sampling method were collected from 14 different districts in Guangdong Province[2]. The 14 districts were divided into developed and less developed city groups, with Guangzhou, Foshan, and Shenzhen constituting the developed city group and the rest of the cities being classified as the less developed city groups. After removing the samples obtained from those who did not complete the questionnaire and those with faecal sample sequences of less than 10,000 reads. In addition, a total of 6,376 samples were retained in this study after a preprocessing procedure was performed on the data. We further selected 1083 healthy individuals from the cohort as healthy controls. The inclusion criteria for healthy individuals were as follows: no reported illness; fasting blood glucose (FBG) levels of <6.1; BMI of <24; no antibiotic use within 1 month of stool sample donation; no clinical parameters indicating diagnoses of hypertension, hypertriglyceridaemia, hypercholesterolemia, neurosis, or chronic fatigue syndrome (CFS); and no medication for metabolic diseases intake before donating the stool sample. As the age composition differed significantly between the healthy control group and the different disease groups, we age-matched the different disease groups to the healthy control group using chronological age to avoid its effect as a strong confounding factor. The matched disease groups were as follows: hypertriglyceridaemia (n=537), hypercholesterolemia (n=571), metabolic syndrome (MetS) (n=695), diabetes (n=445), hypertension (n=688) and obesity (n=585).

Sample processing

We described the sample processing method in our previous study[1]. In brief, we extracted total bacterial DNA from stool samples. The barcoded primers (5′ to 3′) by which the 16S rRNA gene V4 region was amplified were V4F, GTGYCAGCMGCCGCGGTAA and V4R, GGACTACNVGGGTWTCTAAT. The PCR amplification conditions were as follows: initial denaturation at 94 °C for 5 min, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 52 °C for 30 s, extension at 72 °C for 45 seconds, and a final elongation at 72 °C for 5 min. Then, the PCR products were sent to Beijing Genome Institute (BGI; Beijing, China), and next-generation sequencing was performed on an Illumina HiSeq 2500 platform.

Bioinformatics and biostatistics

The method used for the preprocessing of the raw sequences was described in our previous study[1]. Briefly, using QIIME V.1.9.1, we merged the data, performed quality control assessments and analysed the data at the sequence level using the reference-free statistical denoising deblur method to prevent improperly identifying distinct operational taxonomic unit-based data originating from misclustered sequences[2, 3].

To establish a microbiome model to predict the age of the host's microbiota, we adopted the relative abundance data at the ASV level of 1083 relatively healthy people to train a random forest regression model by using the R packages ‘randomForest’ (version 4.6.12) and ‘caret’ (version 6.0.68) to tune hyperparameters and perform 10-fold cross-validation. The sample matching process was performed using the R package 'MatchIt'. It used the propensity score method (PSM) and greedy nearest neighbour matching to achieve pairwise matching of participants in different disease groups with their healthy control counterparts. Spearman's rank correlation test was used to analyse the correlation between amplicon sequence variant (ASVs) and host chronological age, and then $P \leq 0.05$ was used as the threshold to filter out the ASVs related to chronological age for modeling. The interaction effect between chronological age and different metabolic disorders was indicated by the significance of the interaction term in the multivariable linear regression model (Relative abundance of the ASV or the microbiota age ~ chronological age + disease group + chronological age*disease group). All plots in this study were drawn using the R package ‘ggplot2’ and ‘pheatmap’, and the statistical indicators in the plots were calculated using the R packages ‘ggpubr’ and ‘ggpmisc’.

Ethics approval

Ethical approvals were described previously[1]. The Guangdong gut microbiome project was approved by the ethical review committee of the Chinese Center for Disease Control and Prevention (No. 201519-A). Written consent was obtained from all participants.

Patient and public involvement

Neither patients nor the public were involved in the design, conduction, reporting, or dissemination plans of our research.

References for online methods

[1] Y. He, W. Wu, H.-M. Zheng, P. Li, D. McDonald, H.-F. Sheng, M.-X. Chen, Z.-H. Chen, G.-Y. Ji, Z.-D.-X. Zheng, P. Mujagond, X.-J. Chen, Z.-H. Rong, P. Chen, L.-Y. Lyu, X. Wang, C.-B. Wu, N. Yu, Y.-J. Xu, J. Yin, J. Raes, R. Knight, W.-J. Ma, and H.-W. Zhou, Regional variation limits applications of healthy gut microbiome reference ranges and disease models. Nature Medicine, 2018. 24(10): 1532-1535.

[2] Lao X Q , Ma W J , Sobko T , et al. Overall obesity is leveling-off while abdominal obesity continues to rise in a Chinese population experiencing rapid economic development: analysis of serial cross-sectional health survey data 2002–2010[J]. International Journal of Obesity.

[3] J.G. Caporaso, J. Kuczynski, J. Stombaugh, K. Bittinger, F.D. Bushman, E.K. Costello, N. Fierer, A.G. Pena, J.K. Goodrich, J.I. Gordon, G.A. Huttley, S.T. Kelley, D. Knights, J.E. Koenig, R.E. Ley, C.A. Lozupone, D. McDonald, B.D. Muegge, M. Pirrung, J. Reeder, J.R. Sevinsky, P.J. Tumbaugh, W.A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, and R. Knight, QIIME allows analysis of high-throughput community sequencing data. Nature Methods, 2010. 7(5): 335-336.

[4] A. Amir, D. McDonald, J.A. Navas-Molina, E. Kopylova, J.T. Morton, Z. Zech Xu, E.P. Kightley, L.R. Thompson, E.R. Hyde, A. Gonzalez, and R. Knight, Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. mSystems, 2017. 2(2).

Supplementary tables

Table S1. Characteristics of the participants included in this study.

Characteristic	**Healthy**, N = 1,083	**Diabetes**, N = 555	**Hypercholesterolemia**, N = 775	**Hypertension**, N = 2,289	**Hypertriglyceridemia**, N = 674	**MetS**, N = 1,254	**Obesity**, N = 633	**p-value**[1]
Age, Median (IQR)	43 (31 – 54)	59 (52 – 66)	57 (49 – 65)	62 (53 – 70)	54 (45 – 63)	58 (49 – 66)	53 (44 – 62)	<0.001
Gender(Female), n (%)	638 (59)	298 (54)	431 (56)	1,194 (52)	297 (44)	601 (48)	379 (60)	<0.001
Weight(kg), Median (IQR)	53 (48 – 58)	61 (54 – 68)	58 (51 – 66)	59 (52 – 67)	64 (57 – 72)	66 (59 – 74)	74 (68 – 81)	<0.001
BMI, Median (IQR)	21.2 (19.6 – 22.5)	24.4 (22.3 – 27.1)	23.4 (21.3 – 25.9)	24.1 (21.8 – 26.5)	25.1 (22.9 – 27.6)	26.1 (24.1 – 28.4)	29.4 (28.6 – 31.2)	<0.001
Waist(cm), Median (IQR)	73 (68 – 78)	85 (79 – 92)	81 (75 – 88)	84 (77 – 90)	86 (80 – 92)	90 (85 – 95)	95 (90 – 100)	<0.001
FBG(mmol/L), Median (IQR)	5.04 (4.68 – 5.34)	8.06 (7.19 – 10.32)	5.47 (5.05 – 6.06)	5.52 (5.08 – 6.14)	5.74 (5.28 – 6.50)	6.11 (5.35 – 6.98)	5.52 (5.07 – 6.28)	<0.001
TCHO(mmol/L), Median (IQR)	5.00 (4.52 – 5.46)	5.39 (4.92 – 5.93)	6.56 (6.35 – 6.86)	5.36 (4.85 – 5.90)	5.40 (4.89 – 5.99)	5.32 (4.80 – 5.89)	5.28 (4.80 – 5.82)	<0.001
TG(mmol/L), Median (IQR)	0.81 (0.64 – 1.10)	1.53 (1.09 – 2.32)	1.30 (0.91 – 1.96)	1.22 (0.85 – 1.80)	3.15 (2.62 – 4.10)	2.06 (1.53 – 2.94)	1.48 (1.05 – 2.23)	<0.001
UA(μmol/L), Median (IQR)	304 (258 – 359)	322 (269 – 401)	329 (273 – 398)	344 (284 – 412)	390 (312 – 459)	368 (309 – 442)	365 (311 – 438)	<0.001
HDL(mmol/L), Median (IQR)	1.31 (1.10 – 1.56)	1.07 (0.89 – 1.34)	1.33 (1.13 – 1.59)	1.21 (0.99 – 1.48)	0.89 (0.73 – 1.08)	0.94 (0.81 – 1.08)	1.07 (0.87 – 1.29)	<0.001
LDL(mmol/L), Median (IQR)	2.88 (2.40 – 3.37)	3.39 (2.75 – 4.10)	4.31 (3.65 – 4.95)	3.41 (2.80 – 4.02)	3.53 (2.87 – 4.23)	3.43 (2.84 – 4.10)	3.49 (2.81 – 4.08)	<0.001
HbA1c(%), Median (IQR)	4.70 (4.30 – 5.10)	6.10 (5.20 – 7.60)	5.10 (4.60 – 5.50)	5.00 (4.60 – 5.50)	5.00 (4.60 – 5.60)	5.20 (4.70 – 5.80)	5.10 (4.60 – 5.50)	<0.001
Hb(g/L), Median (IQR)	142 (129 – 153)	143 (129 – 156)	144 (131 – 157)	143 (130 – 156)	148 (135 – 162)	147 (133 – 159)	145 (134 – 158)	<0.001
ALT(U/L), Median (IQR)	13 (10 – 18)	17 (12 – 26)	17 (12 – 23)	16 (12 – 23)	20 (14 – 30)	19 (13 – 27)	19 (14 – 29)	<0.001
BUN(mmol/L), Median (IQR)	4.82 (4.04 – 5.68)	5.36 (4.34 – 6.38)	5.42 (4.60 – 6.29)	5.39 (4.52 – 6.38)	5.08 (4.36 – 6.05)	5.13 (4.34 – 6.16)	5.03 (4.20 – 6.03)	<0.001
Antibotics use within 1 month, n (%)	0 (0)	42 (7.9)	48 (6.4)	119 (5.3)	52 (7.9)	79 (6.5)	38 (6.1)	<0.001
Medication use for metabolic diseases, n (%)	0 (0)	184 (34)	126 (17)	542 (24)	141 (22)	328 (27)	131 (21)	<0.001

[1] Kruskal-Wallis rank sum test for continuous variables; Pearson's Chi-squared test for categorical variables

Table S2. The associations of potential covariates with microbiota age in the 1083 healthy participants were measured by the multivariable linear regression model.

Covariate	Forest Plot	Beta[1]	SE[2]	95% CI[2]	p-value
Chronological Age		0.37***	0.030	0.31, 0.43	<0.001
Sleep time per day		0.02	0.028	-0.04, 0.07	0.543
BMI		-0.01	0.028	-0.07, 0.04	0.636
Grains		0.06*	0.029	0.00, 0.11	0.050
Fruits		-0.11***	0.030	-0.16, -0.05	<0.001
Fruit drinks		-0.03	0.031	-0.09, 0.03	0.410
Vegetables		0.01	0.030	-0.05, 0.06	0.839
Livestock meat		-0.01	0.030	-0.07, 0.05	0.654
Carbonated beverage		0.04	0.031	-0.02, 0.10	0.233
High alcohol liquor		0.03	0.028	-0.02, 0.09	0.237
Low alcohol liquor		0.00	0.032	-0.07, 0.06	0.922
Beer		-0.03	0.031	-0.09, 0.03	0.337
Yellow rice wine		-0.03	0.028	-0.09, 0.02	0.251
Rice wine		0.07*	0.032	0.01, 0.13	0.028
Wine		-0.05	0.031	-0.11, 0.01	0.126

[1] Beta is the standardized regression coefficient in the linear regression model.
*p<0.05; **p<0.01; ***p<0.001

[2] SE = Standard Error, CI = Confidence Interval

Supplementary figures

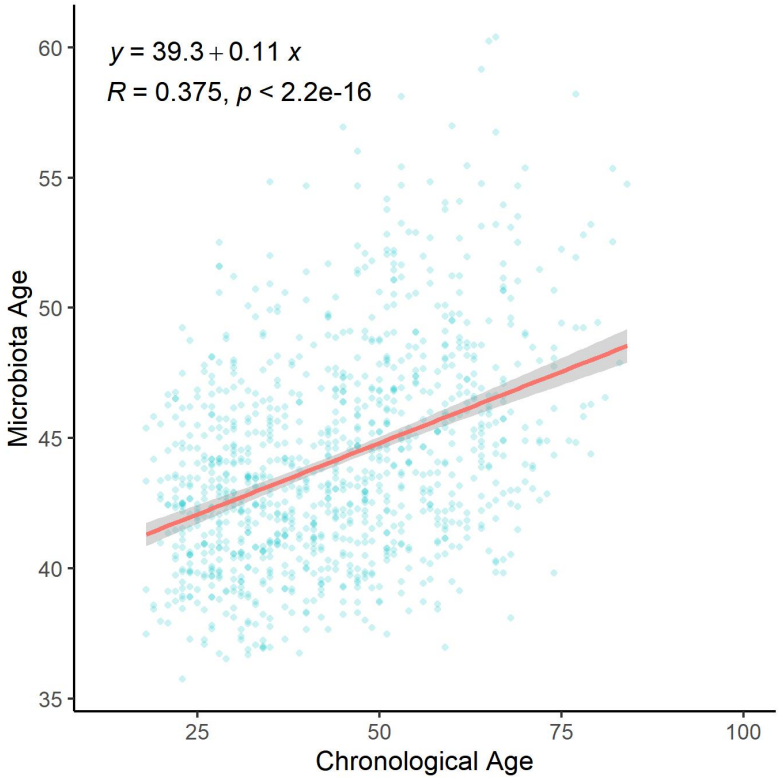

Figure S1. The scatter plot showed a significant and positive correlation between the microbiota age of 1083 healthy individuals predicted by the random forest model and their chronological age. R, Spearman correlation coefficient.

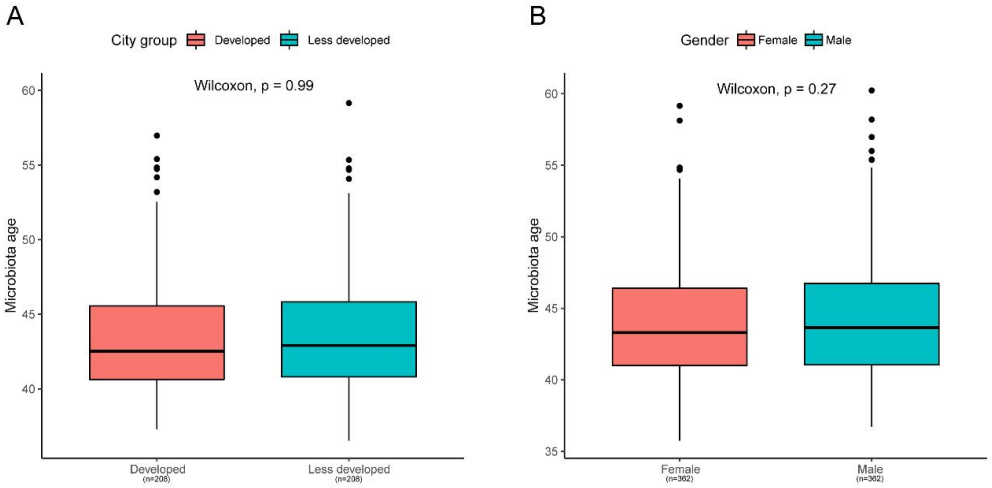

Figure S2. Comparing gut microbiota age between geographic areas (A) and gender (B) in the healthy individuals. Chronological ages were matched between different groups using the propensity score method.

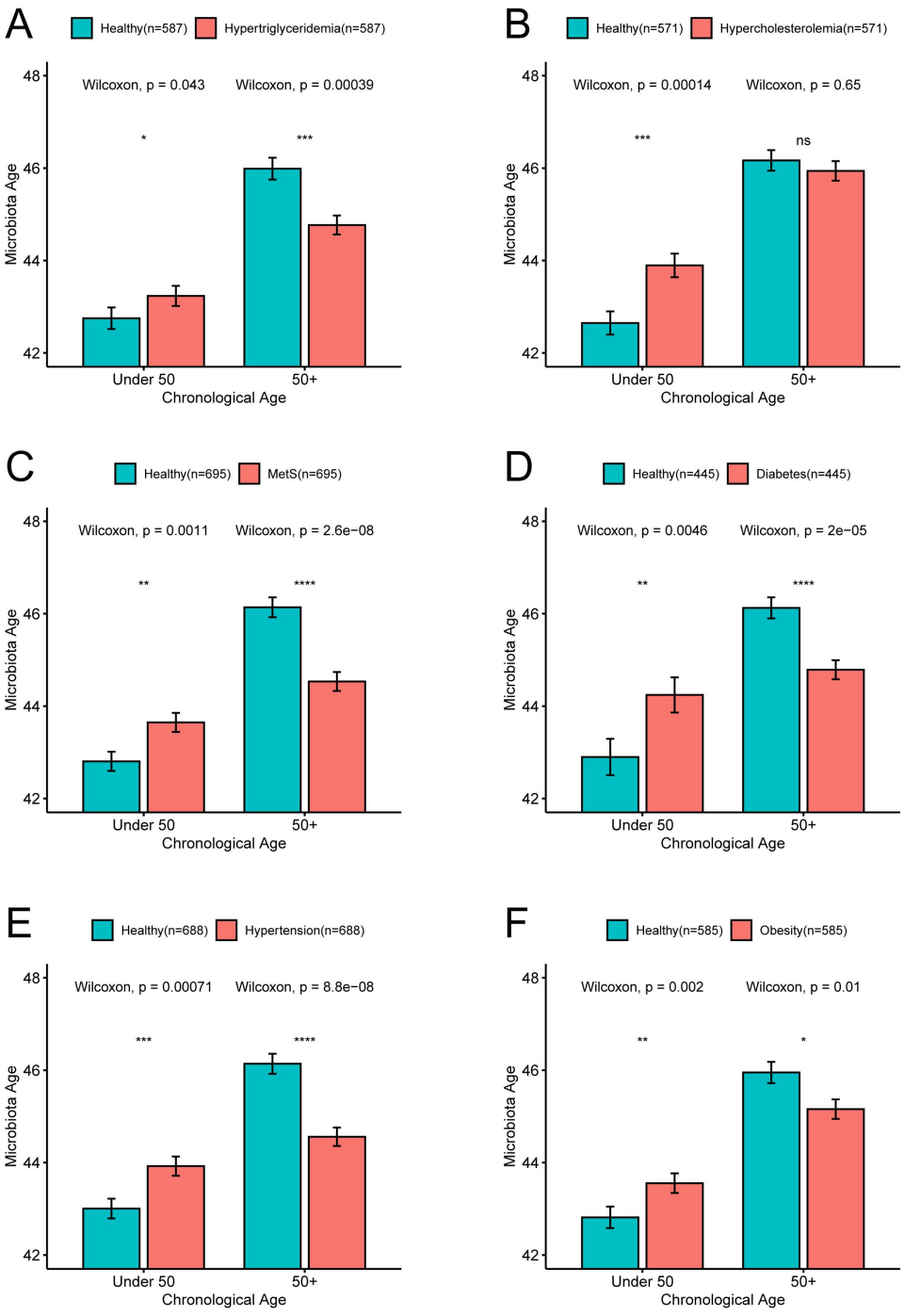

Figure S3. Comparing microbiota age between people with different metabolic disorders and healthy individuals. Figures S2(A) to S2(F) illustrate the comparisons between groups with hypertriglyceridaemia, hypercholesterolemia, metabolic syndrome (MetS), diabetes, hypertension, obesity and healthy individuals, respectively. The "under50" and "50+" labels indicate that the chronological age of the participant in their corresponding group is less than 50 years and greater than or

equal to 50 years, respectively. Significantly different groups are indicated with **** for $P \leqslant 0.0001$, *** for $P \leqslant 0.001$, ** for $P \leqslant 0.01$, * for $P \leqslant 0.05$ or ns for $P > 0.05$ (two-sided Wilcoxon test).

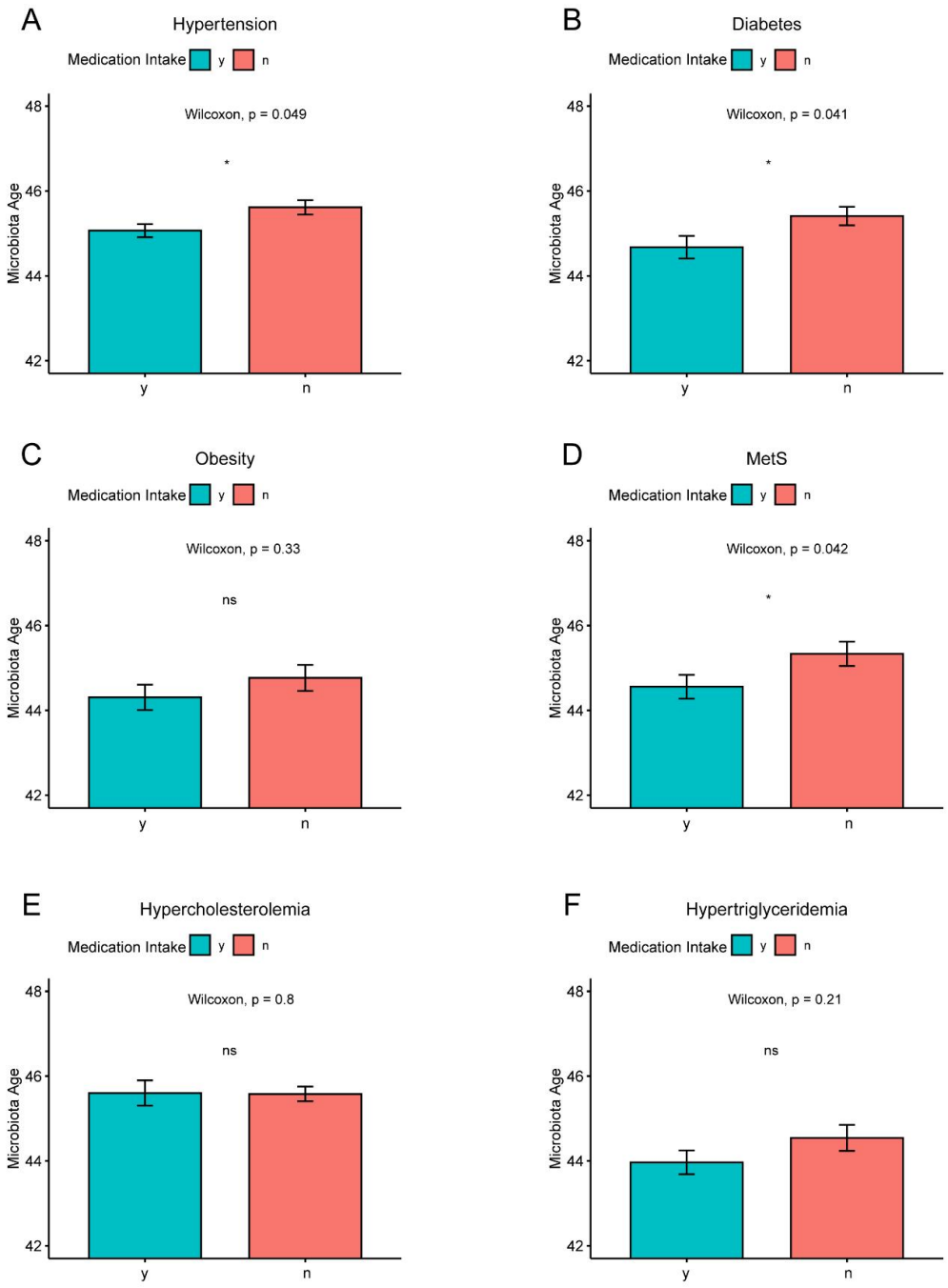

Figure S4. Comparison of gut microbiota age between metabolic diseases patients who took medication and those who are unmedicated. Figures S3(A) to S3(F) show the differences in microbiota age of individuals with hypertension, diabetes, obesity, metabolic syndrome (MetS), hypercholesterolemia and hypertriglyceridemia respectively after matching for chronological age between the two groups. Significantly different groups are indicated with **** for P⩽ 0.0001, *** for P ⩽ 0.001, ** for P ⩽ 0.01, * for P ⩽ 0.05 or ns for P > 0.05 (two-sided Wilcoxon test).

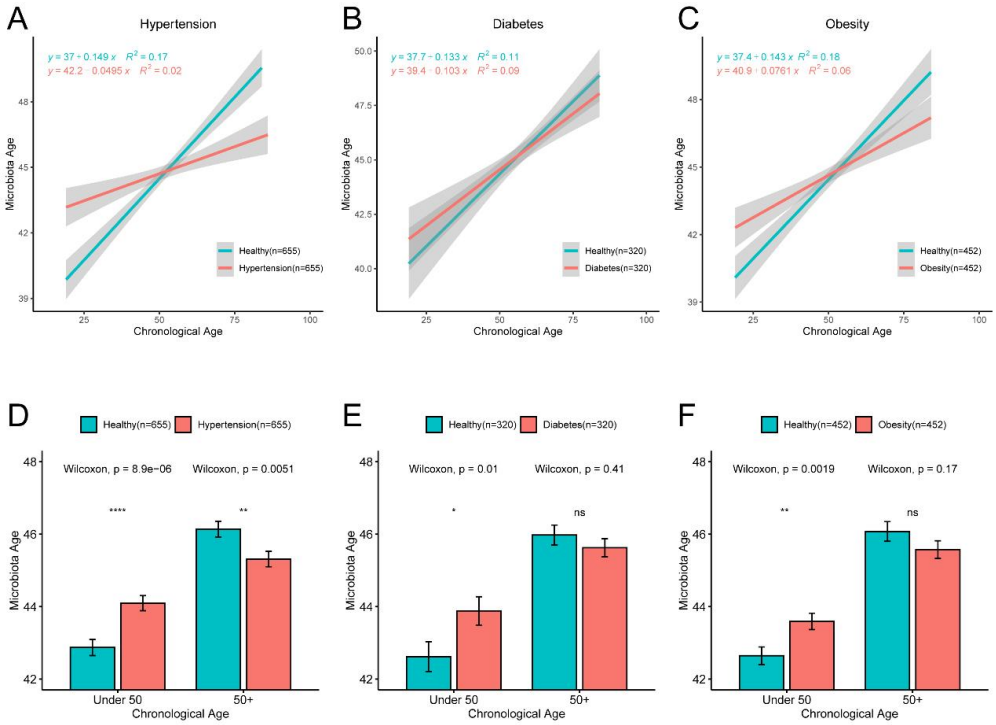

Figure S5. Reproducing the intersecting pattern between gut microbiota age and metabolic diseases after excluding medicated patients. In order to avoid the effect of antibiotic use and intake of medication on our results, we excluded subjects who used antibiotics within one month and used medication for metabolic diseases before sample collection in different disease groups. Meanwhile, patients' chronological age was matched with healthy individuals. Figures S3(A) to S3(C) illustrate the microbiota ageing trajectories of chronological age-matched participants in hypertension, diabetes and obesity respectively. Significantly different groups are indicated with **** for P⩽ 0.0001, *** for P ⩽ 0.001, ** for P ⩽ 0.01, * for P ⩽ 0.05 or ns for P > 0.05 (two-sided Wilcoxon test).

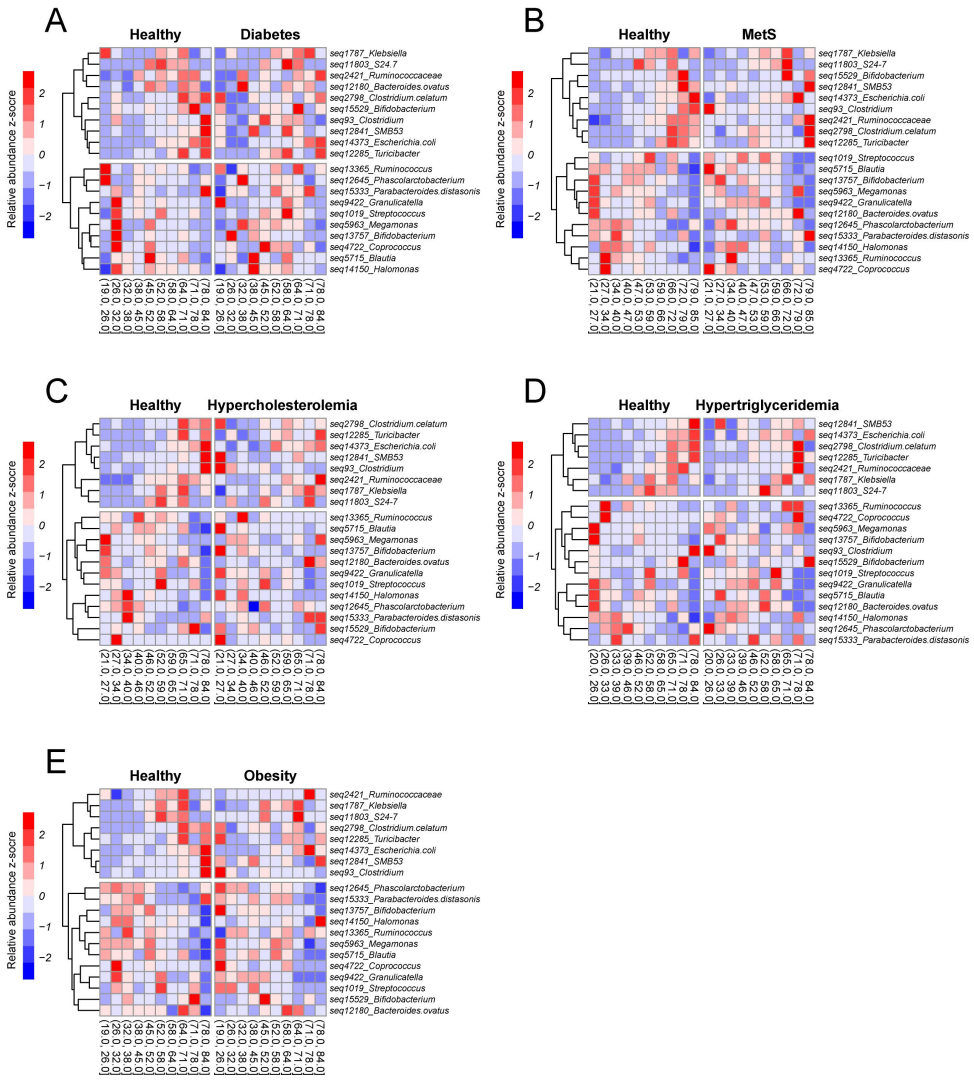

Figure S6. The aging trajectories of gut microbes in participants with different metabolic diseases and healthy individuals. The healthy and diseased groups were matched by chronological ages.