

University of Groningen

The long-term genetic stability and individual specificity of the human gut microbiome

Lifelines Cohort Study; Chen, Lianmin; Wang, Daoming; Garmaeva, Sanzhima; Kurilshikov, Alexander; Vich Vila, Arnau; Gacesa, Ranko; Sinha, Trishla; Segal, Eran; Weersma, Rinse K

Published in:
Cell

DOI:
[10.1016/j.cell.2021.03.024](https://doi.org/10.1016/j.cell.2021.03.024)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2021

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Lifelines Cohort Study, Chen, L., Wang, D., Garmaeva, S., Kurilshikov, A., Vich Vila, A., Gacesa, R., Sinha, T., Segal, E., Weersma, R. K., Wijmenga, C., Zhernakova, A., & Fu, J. (2021). The long-term genetic stability and individual specificity of the human gut microbiome. *Cell*, 184(9), 2302-+. <https://doi.org/10.1016/j.cell.2021.03.024>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

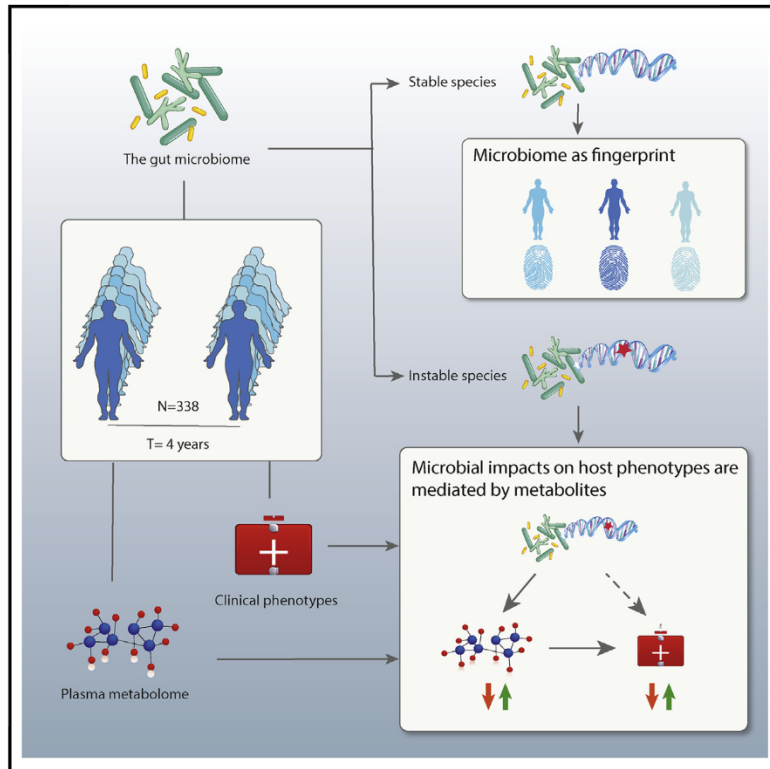
Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

The long-term genetic stability and individual specificity of the human gut microbiome

Graphical abstract



Authors

Lianmin Chen, Daoming Wang, Sanzhima Garmaeva, ..., Cisca Wijmenga, Alexandra Zhernakova, Jingyuan Fu

Correspondence

sashazhernakova@gmail.com (A.Z.), j.fu@umcg.nl (J.F.)

In brief

A longitudinal analysis of 338 individuals across 4 years characterizes the temporal stability and diversity of human gut microbiome, linking the microbial profile with individual-specific host phenotypes and suggesting the role of plasma metabolites in mediating the impact of microbiome on host physiology.

Highlights

- Gut microbial genetic stability differs across species
- Gut microbial composition with higher baseline diversity is more stable over time
- Individual-specific, temporally stable microbial profiles can fingerprint the host
- The microbial impact on host health is mediated by plasma metabolites



Article

The long-term genetic stability and individual specificity of the human gut microbiome

Lianmin Chen,^{1,2,6} Daoming Wang,^{1,6} Sanzhima Garmaeva,¹ Alexander Kurilshikov,¹ Arnau Vich Vila,^{1,3} Ranko Gacesa,^{1,3} Trishla Sinha,¹ Lifelines Cohort Study, Eran Segal,^{4,5} Rinse K. Weersma,³ Cisca Wijmenga,¹ Alexandra Zernakova,^{1,*} and Jingyuan Fu^{1,2,7,*}

¹Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen 9713GZ, the Netherlands

²Department of Pediatrics, University of Groningen, University Medical Center Groningen, Groningen 9713GZ, the Netherlands

³Department of Gastroenterology and Hepatology, Groningen 9713GZ, University of Groningen, University Medical Center Groningen, Groningen 9713GZ, the Netherlands

⁴Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 7610001, Israel

⁵Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot 7610001, Israel

⁶These authors contributed equally

⁷Lead contact

*Correspondence: sashazernakova@gmail.com (A.Z.), j.fu@umcg.nl (J.F.)

<https://doi.org/10.1016/j.cell.2021.03.024>

SUMMARY

By following up the gut microbiome, 51 human phenotypes and plasma levels of 1,183 metabolites in 338 individuals after 4 years, we characterize microbial stability and variation in relation to host physiology. Using these individual-specific and temporally stable microbial profiles, including bacterial SNPs and structural variations, we develop a microbial fingerprinting method that shows up to 85% accuracy in classifying metagenomic samples taken 4 years apart. Application of our fingerprinting method to the independent HMP cohort results in 95% accuracy for samples taken 1 year apart. We further observe temporal changes in the abundance of multiple bacterial species, metabolic pathways, and structural variation, as well as strain replacement. We report 190 longitudinal microbial associations with host phenotypes and 519 associations with plasma metabolites. These associations are enriched for cardiometabolic traits, vitamin B, and uremic toxins. Finally, mediation analysis suggests that the gut microbiome may influence cardiometabolic health through its metabolites.

INTRODUCTION

The human gut harbors a diverse community of microbes that exhibit large between-individual variations (Falony et al., 2016; Lloyd-Price et al., 2017; Rothschild et al., 2018; Zernakova et al., 2016), and cross-sectional analyses have linked these variations to human health and disease phenotypes (Chen et al., 2020a; Falony et al., 2016; Rothschild et al., 2018; Vieira-Silva et al., 2020; Zernakova et al., 2016). The gut microbiota also undergoes compositional changes over the course of an individual's life, as either the cause or consequence of changes in host health and disease status (Chen et al., 2018; Vatanen et al., 2018; Zhou et al., 2019). Several studies have assessed temporal changes in microbial taxonomical composition (Faith et al., 2013; Mehta et al., 2018) and laid the foundation for targeted mechanistic investigations of the consequences of host-microbiome crosstalk for health and disease, including studies in early childhood (Stewart et al., 2018), early-onset type 1 and type 2 diabetes (Vatanen et al., 2018; Zhou et al., 2019), and inflammatory bowel disease (Lloyd-Price et al., 2019).

Nevertheless, several important questions about the temporal variability of the gut microbiome remain unexplored. First, beyond gut microbial composition, the genetic makeup of microbial genomes can also undergo dynamic changes over time. Microbial genomic changes due to evolution and strain replacement, such as single-nucleotide mutations and gain or loss of genomic regions (structural variation), are implicated in the development of human disease (Greenblum et al., 2015; Schloissnig et al., 2013; Zeevi et al., 2019). Yet investigations of temporal changes in microbial genetic makeup are still missing. Second, while cross-sectional association analyses have reported numerous associations with host health and disease (Falony et al., 2016; Lloyd-Price et al., 2017; Rothschild et al., 2018; Zernakova et al., 2016), these associations lack longitudinal confirmation that would allow us to assess whether alterations of the gut microbiome are related to changes in host health status. Third, other microbial components such as antibiotic resistance and virulence factors have become a major concern given the wide-scale use of antibiotics in the last decades. The risk of transfer of resistance and virulence genes between microorganisms has been extensively investigated due to

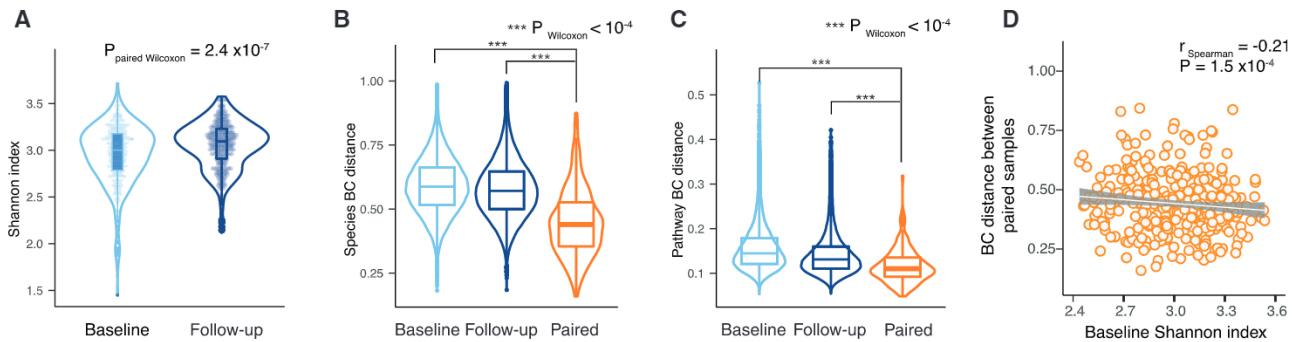


Figure 1. Long-term variation of the gut microbiome composition

(A) The comparison of the microbial alpha diversity at baseline (light blue) and follow-up (dark blue). The y axis refers to the Shannon index at the species level. (B) The comparison of inter-individual Bray-Curtis distance of species composition at baseline (light blue) and follow-up (dark blue), as well as the Bray-Curtis distance within the paired samples between two time points (orange). (C) The comparison of the inter-individual Bray-Curtis distance of functional profile at baseline (light blue) and follow-up (dark blue), as well as the Bray-Curtis distance within paired samples at two time points (orange). (D) Temporal stability of microbial composition might be dependent on its baseline diversity. The x axis refers to baseline Shannon index and the y axis refers to the Bray-Curtis distance within the paired samples between two time points. p values from rank-based Wilcoxon test and the Spearman correlation are shown accordingly.

See also [Figure S2](#).

its relevance to human health (Ochman et al., 2000). However, information on the spread of antibiotic resistance and virulence genes among human gut-commensal microorganisms over time has not been reported, which impedes the effective prevention and treatment of bacterial infections.

In this study, we present a long-term follow-up analysis of the gut microbiome in 338 participants of the population-based Lifelines-DEEP cohort (Tigchelaar et al., 2015) in which we compare samples taken 4 years apart. We characterized long-term temporal stability in gut microbial composition and genetic makeup and aimed to answer two types of questions: (1) which bacterial features are both individual specific and temporally stable? Can we use these features as a “fingerprint” to distinguish samples from the same individual? and (2) which bacterial features show large temporal variation? Can their temporal variation be linked to changes in the host’s clinical phenotypes and lifestyle? To gain further biological insights, we profiled plasma levels of 1,183 metabolites at both time points and used mediation analysis as an *in silico* method to infer whether metabolites mediate the causal relationship behind the microbial impact on host health. Finally, we assessed the changes in antibiotic resistance and virulence factors in the human gut microbiome.

RESULTS

The Lifelines-DEEP follow-up cohort

To investigate the long-term variability of the human gut microbiome, we collected fecal samples from 338 individuals from the prospective, population-based Lifelines-DEEP (LLD) cohort taken 4 years apart (Tigchelaar et al., 2015) and processed these samples using the same lab protocols and bioinformatic pipelines. 51 phenotypic factors were assessed at both time points, including anthropometric traits (e.g., age, sex, and body mass index [BMI]), blood cell counts, biochemical measurements (e.g., glucose, HbA1c, and blood lipid profile), diseases, and medica-

tion usage (Table S1). For 22 continuous traits, we observed significant temporal changes in 19 phenotypic factors at a false discovery rate (FDR)_{Paired Wilcoxon} < 0.05 (Table S1). For instance, significant increases were observed for plasma levels of creatinine ($P_{\text{Paired Wilcoxon}} = 2.5 \times 10^{-50}$) and systolic and diastolic blood pressures ($P_{\text{Paired Wilcoxon}} = 3.6 \times 10^{-26}$ and 2.2×10^{-3}) (Figure S1). We also observed changes in lifestyle, diseases, and medication usage (Table S1). For example, 26 participants developed irritable bowel syndrome (IBS), while 6 developed depression (Table S1).

Temporal changes in gut microbial diversity and composition

To characterize the stability of the gut microbiome over time, we first investigated microbial composition and diversity. Compared to baseline, we observed a significant increase in the alpha-diversity (species-level Shannon index, $P_{\text{Paired Wilcoxon}} = 2.4 \times 10^{-7}$, Figure 1A), as well as a moderate variation in microbial species and pathway abundance ($P_{\text{PCo1 Paired Wilcoxon}} > 0.08$ and $P_{\text{PCo2 Paired Wilcoxon}} < 1.6 \times 10^{-5}$ for both species and pathway, respectively, Figure S2). The differences in overall microbial species and pathway composition were larger between individuals than within individuals ($P_{\text{Wilcoxon}} < 1 \times 10^{-4}$, Figures 1B and 1C), indicating that an individual’s gut microbial composition is more similar to their own composition 4 years ago than to those of other people. Interestingly, within-individual differences in gut microbial species composition were smaller in participants with a higher baseline alpha-diversity ($r_{\text{Spearman}} = -0.21$, $p = 1.5 \times 10^{-4}$, Figure 1D), supporting the hypothesis that diverse microbial communities tend to be more stable (Coyte et al., 2015).

When comparing individual microbial species and pathways, the relative abundance of 59.9% species (94 out of 157) and 44.3% pathways (152 out of 343) showed a significant difference at FDR < 0.05 (paired Wilcoxon test, Table S2). Species belonging to the same genera often showed consistent changes

in direction, e.g., the relative abundance of 7 *Bifidobacterium* species all decreased significantly, while the abundances of the majority of *Alistipes* species (7 out of 8) increased (Table S2). These changes may partially be due to an age effect. For instance, several *Bifidobacterium* species, including *B. adolescentis*, *B. bifidum*, and *B. longum*, have been observed to be negatively associated with age (Zhernakova et al., 2016).

Microbial genetic stability differs substantially across species

Microbial genetic makeup may also change over time, e.g., due to mutagenesis and strain replacement. Characterization of the stable and changeable genetic components of the gut microbiome over a long time course is important for further understanding the importance of microbial strain alterations with respect to host phenotypic changes. Here, we characterized within-individual temporal microbial genetic differences by comparing both single-nucleotide polymorphism (SNP) haplotypes (Truong et al., 2017) and genomic structural variants (SVs) (Zeevi et al., 2019). SNP haplotype differences were characterized for 37 species that were present in at least 5 paired samples from both time points (Figure 2A; Table S2). We also identified 6,130 SVs, including 4,333 deletion SVs (dSVs, absence of genomic regions) and 1,797 variable SVs (vSVs, genomic regions with variable coverage) from 41 microbial species present in at least 5 paired samples (Figure 2B; Table S2).

We observed that within-individual genetic differences in SNP haplotypes and SVs were significantly smaller than the differences between different individuals (Figures 2A and 2B; Table S2). The species that showed large temporal changes in their SNP haplotypes included *Ruminococcus torques*, *Streptococcus parasanguinis*, and *Faecalibacterium prausnitzii*, while *Bifidobacterium angulatum*, *Methanobrevibacter smithii*, and *Alistipes putredinis* showed relatively low genetic variability ($P_{\text{Wilcoxon}} < 0.05$, Figure 2A). A consistent trend in the genetic stability of SNP profiles was also observed in 43 healthy participants with fecal microbiome data available 1 year apart from the Human Microbiome Project (HMP) (Figure 2C) (Schloissnig et al., 2013). Compared to the HMP cohort, the genetic differences in unstable species were larger in the LLD cohort, potentially due to a longer duration of follow-up (Figures 2D–2F). This observation further supports the genetic instability of these species over time.

For 23 species, both strain SNP haplotype and SV information were available (Figures 2A and 2B). Temporal variability in SNP haplotypes and SVs showed substantial consistency (Figures S3A and S3B), illustrating that the microbial genetic stability of some species can be seen at different levels of genetic variation. For example, several species with highly time-dependent variability in SNP haplotypes, such as *R. torques* and *F. prausnitzii*, also showed a high degree of changes in their SVs, whereas other species, such as *M. smithii*, showed high stability of both SNP haplotypes and within-individual SV variability (Figures 2A and 2B).

Interestingly, these genetically unstable species have often been reported to be related to human health and disease. For instance, previous studies have found a higher abundance of

R. torques in patients with Crohn disease (Joossens et al., 2011), a higher level of *S. parasanguinis* in patients with intestinal infection (Vacca, 2017) and a lower level of *F. prausnitzii* in patients with inflammatory bowel disease (Munukka et al., 2017; Vich Vila et al., 2018). Notably, within-individual differences in microbial genetic makeup did not correlate with changes in abundance (Figure S3C), suggesting that microbial genetic variability provides an extra layer of information that is independent of microbial abundance. Furthermore, we hypothesized that the temporal changes in genetic makeup can be also driven by replacement of the dominant strain. Using SNP profiles, we observed distinct strains for five species, *R. torques*, *F. prausnitzii*, *S. parasanguinis*, *Ruminococcus obeum*, and *Eubacterium rectale*, that showed >70% genetic dissimilarity (Figure S4).

Taken together, these results illustrate that within-individual differences in both microbial composition and genomes can be detected 4 years apart, but within-individual similarity of microbiome compositional and genetic profiles is greater than between-individual similarity. The stable and variable microbial compositional and genetic components we observe can have different implications: individually stable microbial components might be used to identify their host, whereas variable microbial components might be related to phenotypic changes of the host.

Microbial genetic makeup shows individuality that can serve as a host fingerprint

We observed that some species, such as *M. smithii*, showed large between-individual variability but small within-individual differences in their genetic makeup (Figure 2A). Per 100 base pairs (bp) of the species-specific regions, *M. smithii* had an average 0.11 bp difference between two samples from the same individual but an average 2.77 bp difference between different individuals ($P_{\text{Wilcoxon test}} = 3.6 \times 10^{-64}$, Figure 2A; Table S2). This inspired us to evaluate the possibility of using microbial genetic and compositional profiles to distinguish samples from the same individuals. We generated the SNP haplotype profiles of *M. smithii* for 100 paired samples. Based on the distance of the *M. smithii* SNP profiles, we could correctly link 94 paired samples, resulting in an accuracy of 94% (proportion of correctly distinguished sample pairs) (Figure S5). We systematically assessed the fingerprinting potential of the 71 microbial profiles that were present in >10% of paired samples (Table S3), including the composition of microbial species and metabolic pathways, SNP profiles in 18 species, vSV profiles in 25 species, and dSV profiles in 26 species. In total, 17 of these profiles had at least 50% accuracy in classifying samples. After *M. smithii*, the second-best classifier was the SNP profile of *Phascolarctobacterium succinatutens*, which classified 41 paired samples with 88% accuracy (Figure S5; Table S3). However, sample classification based on microbial species and pathway composition resulted in only 12% and 5% accuracy, respectively (Table S3). Our data imply that microbial genetic profiles dominate over microbial abundance in fingerprinting the human host.

SNP haplotype profiles like *M. smithii* could only be generated for 100 of the 338 paired samples, potentially due to microbial

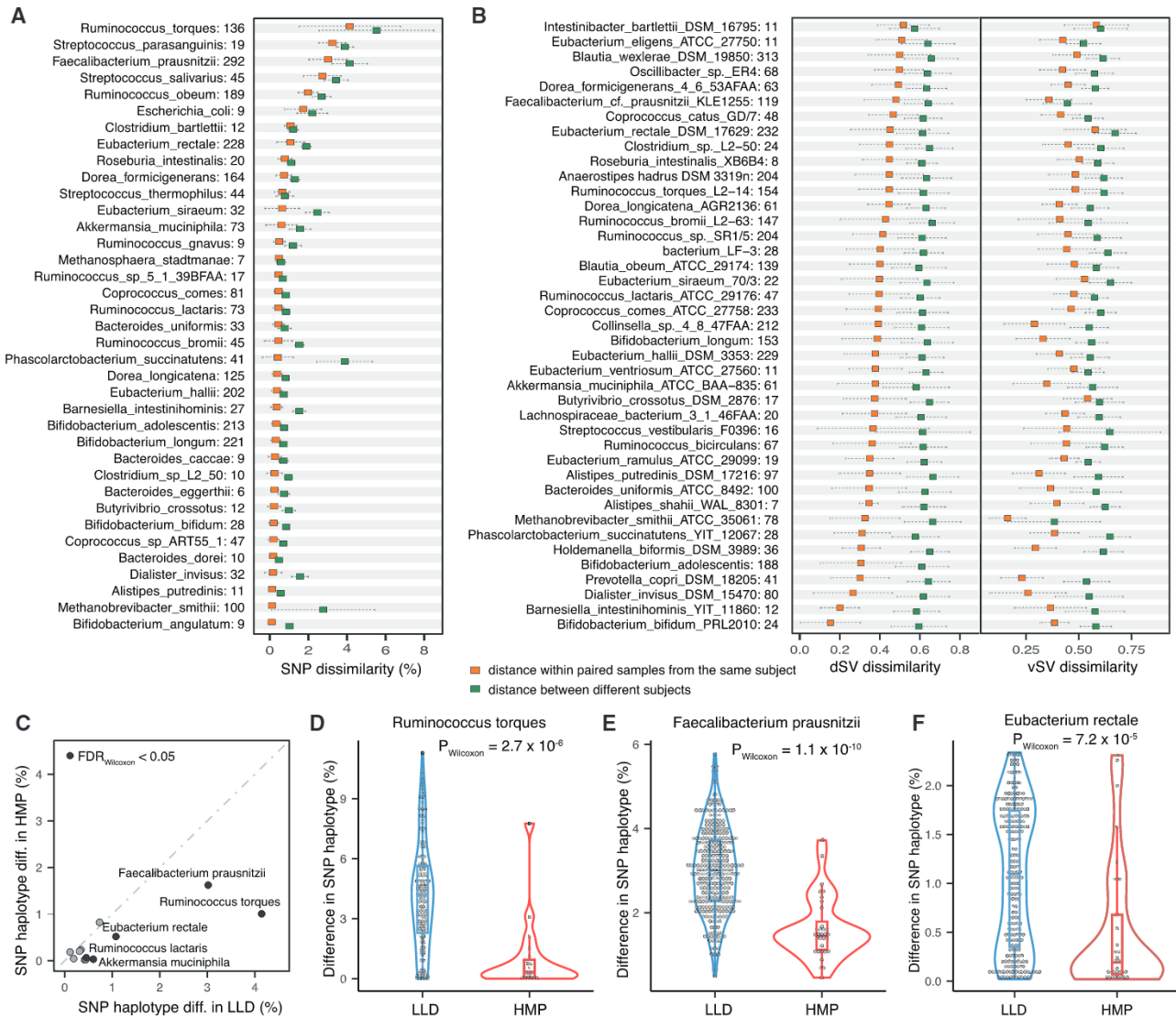


Figure 2. Long-term stability of microbial species SNP haplotypes and structural variants

(A) Within paired samples (orange) and between-individual (green) differences in the single-nucleotide polymorphism (SNP) haplotypes of dominant strains of microbial species. Numbers following species names indicate the number of paired samples for which SNP haplotype profiles are available.

(B) Within paired samples (orange) and between-individual (green) difference in the deletion (dSV) and variable structural variants (vSVs) of microbial strains. Numbers following species names indicate the number of paired samples for which SVs profiles are available.

(C) Comparison of within-individual microbial species SNP haplotype differences between the LLD (4 years apart) and the HMP (1 year apart) ($r_{\text{Pearson}} = 0.82$, $p = 0.001$). Each dot represents one species. Dark dots represent species whose degree of temporal changes is significantly different between LLD and HMP at FDR < 0.05 (Wilcoxon test). Light-gray dots represent the species that showed a similar degree of temporal changes in both cohorts.

(D–F) Comparison of temporal stability of SNP haplotypes of *Ruminococcus torques* (D), *Faecalibacterium prausnitzii* (E), and *Eubacterium rectale* (F) between LLD (in blue) and HMP (in red). The p values from the Wilcoxon tests are shown, respectively.

See also [Figure S3](#) and [Table S2](#).

absence in the gut, which prohibits the use of the genetic profile of a single microbial profile as a host fingerprint. This limitation further inspired us to combine multiple microbial genetics and composition (both species and pathway abundances) for a broader application. We applied stepwise forward selection to optimize the combination of different profiles in a randomly selected 60% of individuals and validated the individual recogni-

tion abilities in the remaining 40% of individuals. The resampling and feature selection were repeated 10 times ([Figures S6A](#) and [S6B](#)). The combination of all 71 profiles ([Table S3](#)) resulted in up to 85% classification accuracy ([Figure 3A](#)), with an optimal model combining the top 30 profiles yielding 82% classification accuracy ([Figure 3B](#)). This optimal model includes SNP profiles of 13 species, dSV profiles of 11 species, vSV profiles of

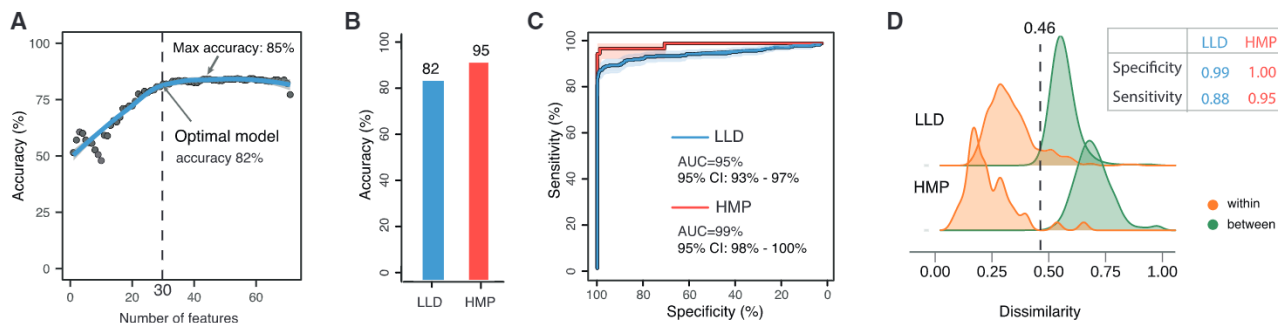


Figure 3. Performance of the gut microbiome in fingerprinting its human host

(A) The combination of all microbial genetic and compositional profiles resulted in up to 85% accuracy in classifying longitudinal samples of 338 individuals in the LLD cohort. The optimal fingerprinting model with 30 microbial features resulted in an accuracy of 82% in the LLD. The x axis refers to the number of features that were included in the stepwise feature selection. The y axis refers to the classification accuracy. Gray dots refer to the accuracy of the stepwise models. The blue line refers to the fitted line of accuracy across all models.

(B) Comparison of the fingerprinting accuracy of the optimal model in the LLD (in blue) and in the HMP (in red) cohorts.

(C) The ROC analysis to show the performance of the fingerprint model in the LLD (in blue) and the HMP (in red) cohorts, respectively. The area under the curve and 95% confidence intervals of both cohorts are shown.

(D) The distribution of within- and between-individual distances of the optimal model. In the LLD cohort, the optimal classification performance in terms of both specificity (0.99) and sensitivity (0.88) was reached at a distance cutoff of 0.46. In the HMP cohort, the specificity and sensitivity were 1 and 0.95, respectively, at the same cutoff.

See also [Figures S5](#) and [S6](#) and [Table S3](#).

5 species, and the Bray-Curtis dissimilarity of species abundance ([Figure S6C](#)). We conducted specificity and sensitivity analyses for the sample classification. The total area under the curve (AUC) was 95% ([Figure 3C](#)), and we reached the optimal 99% specificity and 88% sensitivity at a distance cutoff of 0.46 ([Figure 3D](#)). At this cutoff, we could classify 298 paired samples with 93% accuracy.

We then tested our microbial fingerprinting method in the longitudinal sample collection of 43 individuals in the HMP cohort, our model resulted in 100% accuracy for 41 out of the 43 paired samples at a distance cutoff 0.46 ([Figure 3B](#)) and 95% accuracy in the total set of 43 pairs. This accuracy is higher than the previously reported 30% accuracy based on microbial species abundance and the 80% accuracy based on the microbial gene abundance ([Franzosa et al., 2015](#)). The HMP result confirms the robustness of our microbial fingerprinting method.

Microbial abundance and genomic changes associated with host phenotypes

To examine the role of the gut microbiota in host health, we explored the associations between microbial compositional and genomic changes and host phenotypic changes. We performed two-step analyses to reveal microbial associations to host phenotypes using longitudinal data. First, we performed joint association analyses between microbial features and 51 host phenotypic factors that were highly prevalent between the two time points ([Table S1](#)) using mixed models and including age, sex, and sampling time as covariates. Next, for associations identified at $FDR < 0.05$, we conducted regression analysis on temporal differences, i.e., associations between microbial changes and host phenotypic changes over time (delta association). This identified 190 associations involving 169 microbial features and 34 phenotypes that were significant at $FDR < 0.05$ in

the joint association analysis and also significant at $p < 0.05$ for the delta association analysis with a consistent direction of effect. These included 84 associations with species and pathway abundances and 106 associations with microbial SVs ([Table S4](#)). BMI, blood pressure, HbA1C, and depression were the phenotypic factors with the most associations ([Figure 4A](#)). For instance, we observed a positive association between systolic blood pressure and the abundance of *Lachnospiraceae bacterium* ($\beta_{\text{delta}} = 0.24$, $P_{\text{delta}} = 1.1 \times 10^{-5}$, [Figure 4B](#)) and a negative association between glycosylated hemoglobin (HbA1c) and the flavin biosynthesis pathway (PWY-6168, $\beta_{\text{delta}} = -0.22$, $P_{\text{delta}} = 4.9 \times 10^{-5}$, [Figure 4C](#)). In addition, temporal changes in microbial SVs were associated with host immune phenotypes. For instance, a vSV (3,019–3,020 kb) in *Blautia obeum* that contains virulence protein E and chloramphenicol resistance genes was negatively associated with a change in blood lymphocyte cell counts ($\beta_{\text{delta}} = -0.29$, $P_{\text{delta}} = 6.5 \times 10^{-4}$, [Figure 4D](#)). For disease onset, we detected 22 associations, namely, 15 associations for depression, 3 associations for IBS, and 3 for asthma. The top association for depression onset was found for a dSV region (from 1,164 to 1,165 kb) in *Collinsella* sp. that encodes the histidine kinase A ($P_{\text{Fisher exact test}} = 3.0 \times 10^{-3}$, [Figure 4E](#)).

Microbial abundance and genomic changes are associated with plasma metabolites

To further understand the potential mechanisms by which the gut microbiota could drive host pathophysiology, we hypothesized that metabolites are an important class of molecules that are involved in the host-microbe interaction. By profiling plasma levels of 1,183 metabolites at both time points using untargeted LC-MS, we observed that 27% of metabolites showed significant differences between the two time points at $FDR < 0.05$ (paired Wilcoxon, [Table S2](#)).

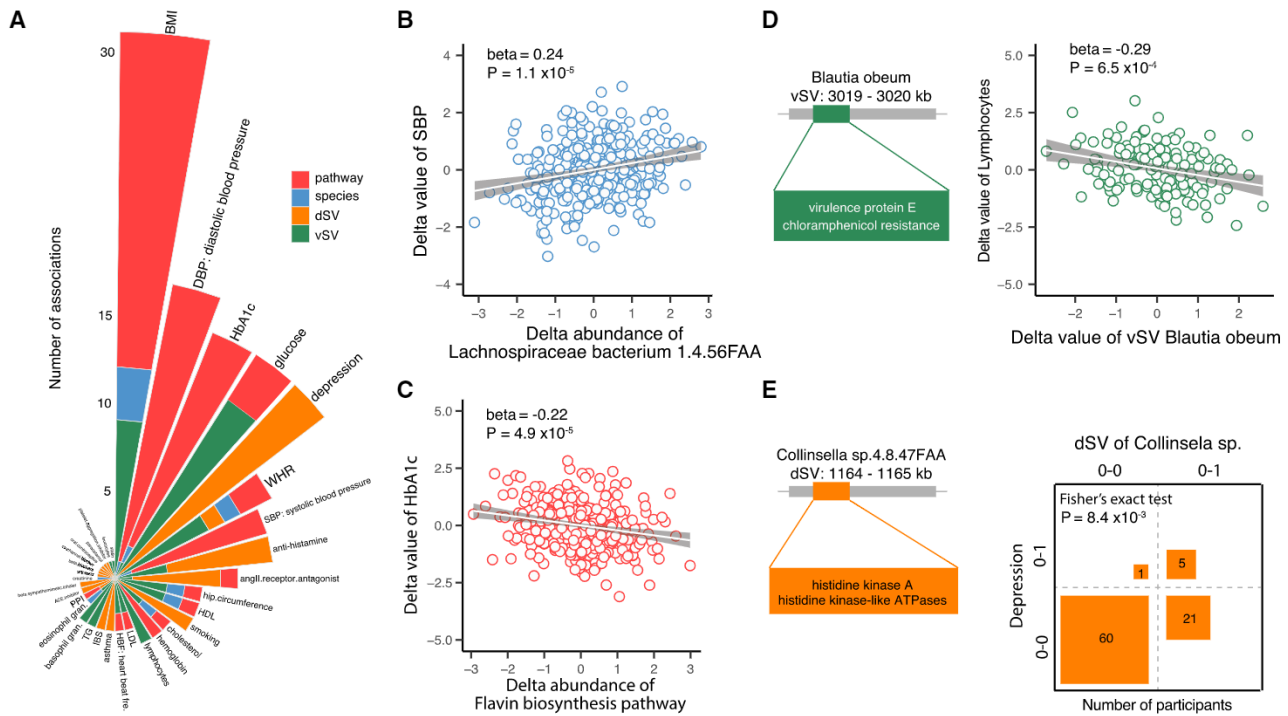


Figure 4. Delta association between phenotype and the gut microbiome

(A) Summary of microbial associations to phenotypes. Each bar represent a phenotype and bars are sorted by the number of associations. The associated microbial factors are colored differently: red for pathway abundance, blue for species abundance, orange for dSVs, and green for vSVs.

(B) Positive delta association between systolic blood pressure and *Lachnospiraceae bacterium*.

(C) Negative delta association between plasma HbA1c and flavin biosynthesis pathway.

(D) Increased *Blautia obeum* vSV (3,019–3,020 kb) associated with the decreased delta value of blood lymphocyte counts. The region of the vSV and the corresponding genes are shown. The delta values refer to the differences of microbial factors and phenotypes between two time points. Beta values refer to the standardized effect size estimated by the linear regression, and the corresponding p values are shown, respectively.

(E) Deletion of *Collinsella sp.* 4.8.47FAA dSV (1,164–1,165 kb) associated with onset of depression. The region of the dSV and the corresponding genes are shown. The association is shown by a 2 × 2 table. The rows refer to disease status, with 0–0 indicating individuals without depression at both time points and 0–1 indicating disease onset. The columns refer to the presence/absence of dSVs of *Collinsella sp.*, with 0–0 indicating individuals with no dSVs at both time points and 0–1 indicating individuals in whom no deletion was detected at baseline but a deletion was detected in the follow-up. The number in each cell indicates the number of individuals per group. The p value from the Fisher's exact test is shown. See also [Figure S1](#) and [Tables S1](#) and [S4](#).

We then looked for metabolic changes specifically in relation to strain replacement and alterations in microbial abundances and SVs. We first checked whether plasma metabolites showed differences in participants with distinct microbial strains ([Figure S5](#)) and, if so, whether strain replacement of these species was related to changes in plasma metabolites. In total, we observed 64 significant associations between 63 metabolites and strain clusters of 5 species ([Table S4](#)). For example, in 292 paired samples, we identified 2 distinct strain clusters of *F. prausnitzii* ([Figure 5A](#)) that associated with 15 metabolites ([Table S4](#)). The top associations were observed for licoriso flavan A, pyrrole, and p-cresol sulfate, and the levels of these metabolites were significantly lower in individuals with *F. prausnitzii* strain 2 ($FDR < 0.05$). We also consistently observed that the abundance of these metabolites decreased in the 24 individuals where *F. prausnitzii* shifted from strain 1 to 2, whereas the metabolite levels increased in the 13 individuals where *F. prausnitzii* shifted from strain 2 to 1 ([Figures 5B–5E](#)). These results suggest that different microbial strains may have different functions that potentially influence host metabolism.

We also detected 455 significant associations between 122 microbial features (species and pathway abundances, dSVs and vSVs) and 81 metabolites ($FDR_{joint} < 0.05$ and $P_{delta} < 0.05$, [Figure 6A](#), [Table S4](#)), including 273 associations with microbial abundance and 182 associations with microbial SVs. Interestingly, various metabolites that associated with the microbiome are already known to be related to the gut microbiome. For instance, we detected 38 microbial associations to plasma levels of thiamine, a vitamin (B1) produced by gut microbes. Thiamine deficiency can affect the cardiovascular system and induce a fast heart rate ([DiNicolantonio et al., 2013](#)). The top microbial associations to thiamine included the species *Alistipes senegalensis* ($\beta_{delta} = 0.20$, $P_{delta} = 4.1 \times 10^{-4}$, [Figure 6B](#)) and *Bacteroidales bacterium* ($\beta_{delta} = 0.23$, $P_{delta} = 5.2 \times 10^{-5}$) and the TCA cycle pathway ($\beta_{delta} = 0.23$, $P_{delta} = 7.2 \times 10^{-5}$) ([Table S4](#)). Notably, the genome of *Alistipes senegalensis* contains genes responsible for thiamine biosynthesis ([Misra et al., 2012](#)).

Another interesting category of metabolites are protein-bound uremic toxins, which are related to microbial metabolism of amino

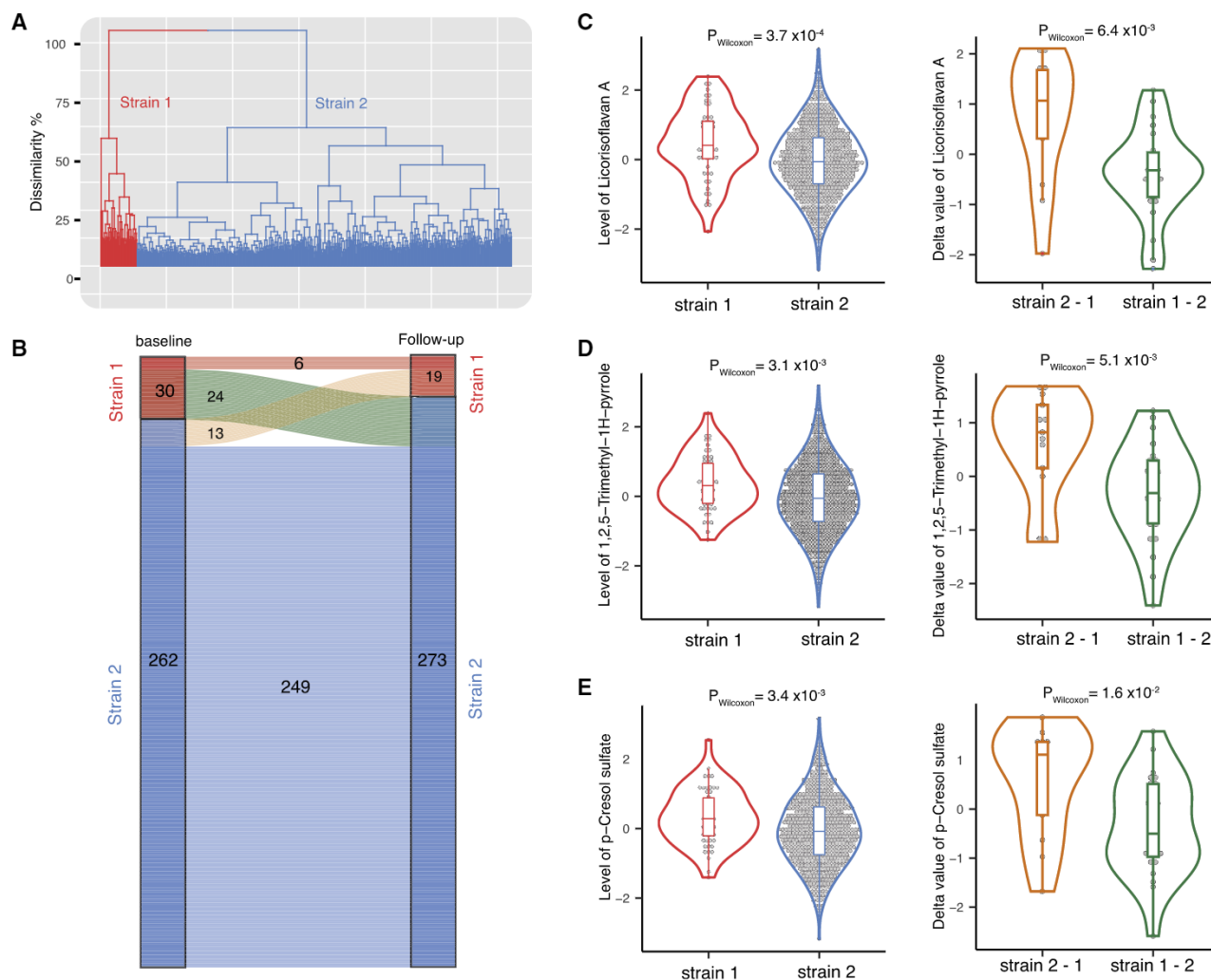


Figure 5. *Faecalibacterium prausnitzii* strain replacement associated with plasma metabolite changes

(A) Two distinct *F. prausnitzii* strain clusters based on its SNP haplotype profile, which are colored in red for strain 1 and blue for strain 2.

(B) Strain replacement of *F. prausnitzii*. Individuals who remained strain 1 and those who remained strain 2 are colored red and blue, respectively. Individuals who switched from strain 1 to 2 are colored green. Individuals who switched from strain 2 to 1 are colored orange. The number of individuals is shown in the plot accordingly.

(C–E) Association of strain replacement with delta value of plasma levels of (C) licoriso flavan A, (D) 1,2,5-Trimethyl-1H-pyrrole, and (E) p-cresol sulfate. The left panel shows the joint associations of plasma abundance of metabolites and difference strains. The right panel shows the association of delta value of metabolites with strain replacements. The p value from Wilcoxon test are shown accordingly.

See also [Figure S4](#) and [Table S4](#).

acids and have been associated with various chronic diseases ([Wang and Zhao, 2018](#)). We characterized plasma levels of 58 uremic toxins from metabolite categories of indoxyl sulfate, p-cresol sulfate, phenyl sulfate, phenylacetic acid, and hippuric acid ([Wang and Zhao, 2018](#)) and observed a significant enrichment for microbial associations: 97 associations for 16 uremic toxins ($P_{\text{Fisher's exact test}} = 1.7 \times 10^{-21}$) ([Figure 6A](#); [Table S4](#)). The most-associated uremic toxins included p-cresol (24 associations), p-cresol sulfate (20 associations), and hippuric acid (16 associations) ([Table S4](#)). P-cresol sulfate is a microbiota-derived metabolite that contributes to many biological and biochemical effects, such as albuminuria in diabetic kidney disease ([Kikuchi et al.,](#)

[2019](#)). The top association to p-cresol sulfate was with *Bacteroidales bacterium ph8* ($\beta_{\text{delta}} = 0.21$, $P_{\text{delta}} = 1.9 \times 10^{-4}$, [Figure 6C](#)), a gut microbial species about which little is currently known. Notably, 22.6% (103 out of 455) of the microbial associations with metabolites were related to vSVs of *Blautia wexlerae* ([Figure 6A](#)). Among these, 27 associations were related to different uremic toxins, particularly to hippuric acid ([Figure 6D](#)), an acyl glycine formed from the conjugation of benzoic acid with glycine and associated with phenylketonuria, propionic acidemia, and tyrosinemia ([Duranton et al., 2012](#)). Intriguingly, these vSV regions encode various membranes transporters, amino acid kinases, urease accessory protein, and protein-binding genes ([Table S4](#)).

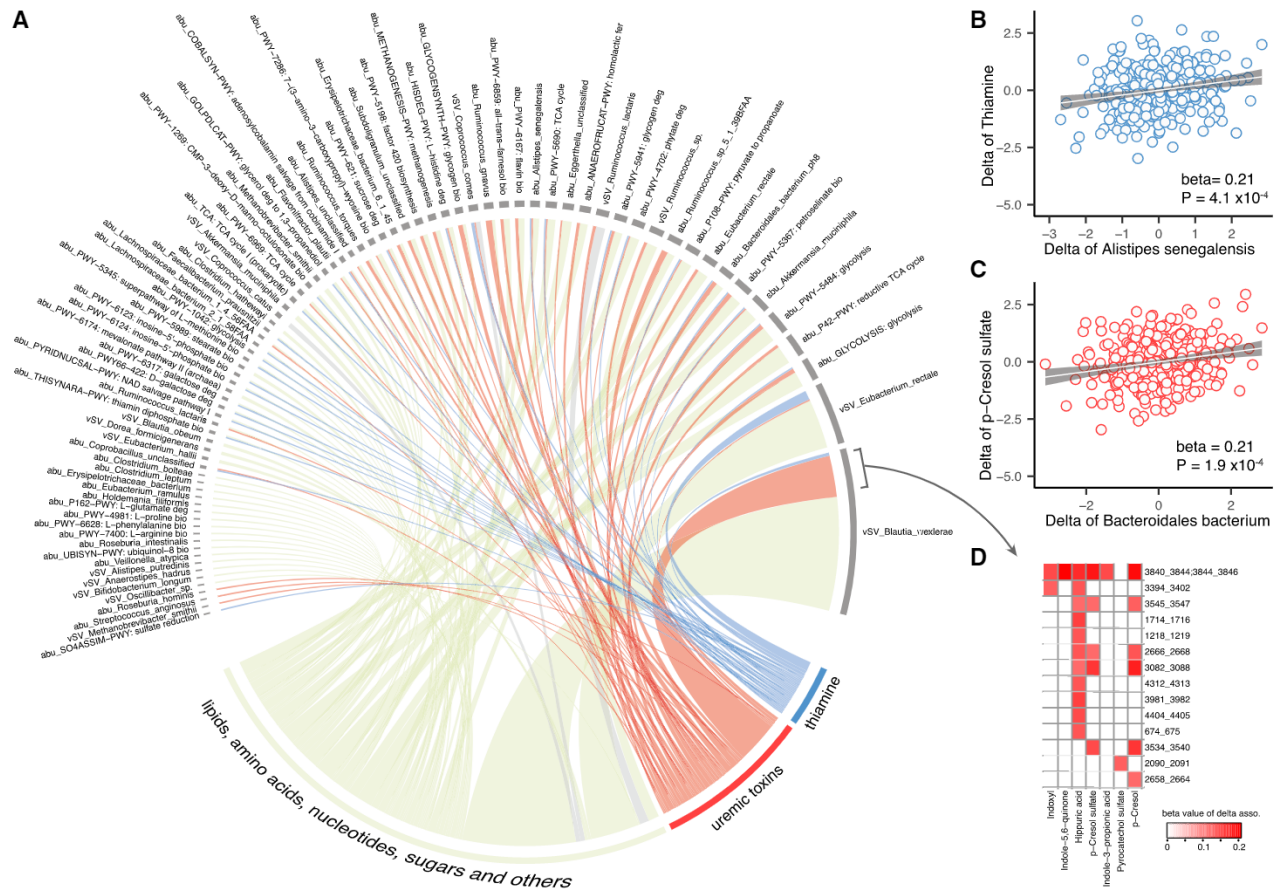


Figure 6. Delta association of plasma metabolite changes with the gut microbiome

(A) Overview of 455 significant microbial associations with plasma metabolites. The associated microbial factors are colored gray, and the associated metabolites are colored blue for thiamine, red for uremic toxins, and green for other metabolites, including lipids, amino acids, nucleotides, and sugars. Each line indicates a significant association between a microbial factor and a metabolite with line color corresponding to the associated metabolite.

(B) A positive delta association between thiamine and *Alistipes senegalensis*. The x axis refers to the delta value of species abundance, i.e., the abundance difference between two time points. The y axis refers to the delta value of thiamine in the plasma. The fitted linear regression line is shown, with standardized beta and p value.

(C) The positive association between microbial-derived uremic toxin p-cresol sulfate and *Bacteroidales bacterium* abundance changes. The x axis refers to the delta value of species abundance, i.e., the abundance difference between two time points. The y axis refers to the delta value of p-Cresol sulfate in the plasma. The fitted linear regression line is shown, with standardized beta and p value.

(D) Heatmap of multiple associations between vSVs of *Blautia wexlerae* and microbial-derived uremic toxins. The cells in red indicate significant associations. The color gradient indicates the standardized beta value from the linear regression, as shown in the color key.

See also Table S4.

The microbiome contributes to host phenotypic changes through metabolites

Of the 169 microbial features associated with clinical phenotypes and the 122 associated with metabolites, 29 microbial features are associated with both (Figure 7A). To evaluate whether metabolites can mediate the microbial impact on host phenotypes, we applied bi-directional mediation analysis, which revealed 21 mediation linkages ($FDR_{\text{mediation}} < 0.05$ and $P_{\text{inverse mediation}} > 0.05$, Figure 7B; Table S4). Most of these linkages were related to microbial impact on blood pressure via thiamine (9 linkages) and acetyl-N-formyl-5-methoxykynurenamine (AFMK, 9 linkages). The impact of thiamine on cardiometabolic health has been well documented and was confirmed in a ran-

domized controlled trial that showed that thiamine can reduce diastolic blood pressure (Alaei-Shahmiri et al., 2015). AFMK is a degradation metabolite of melatonin that contributes to blood pressure reduction by inhibiting the synthesis of prostaglandin (Mayo et al., 2005; Rezzani et al., 2010). Our mediation analysis suggested that various bacterial pathways may contribute to these effects. For instance, the microbial sulfate reduction pathway may contribute to a decrease in diastolic blood pressure by increasing plasma thiamine levels (21%, $P_{\text{mediation}} = 6.0 \times 10^{-3}$, Figure 7C), and bacterial lipopolysaccharide biosynthesis may contribute to a decrease in systolic blood pressure by affecting plasma levels of AFMK (16%, $P_{\text{mediation}} = 6.0 \times 10^{-3}$, Figure 7D). Metabolic products of the bacterial sulfate reduction

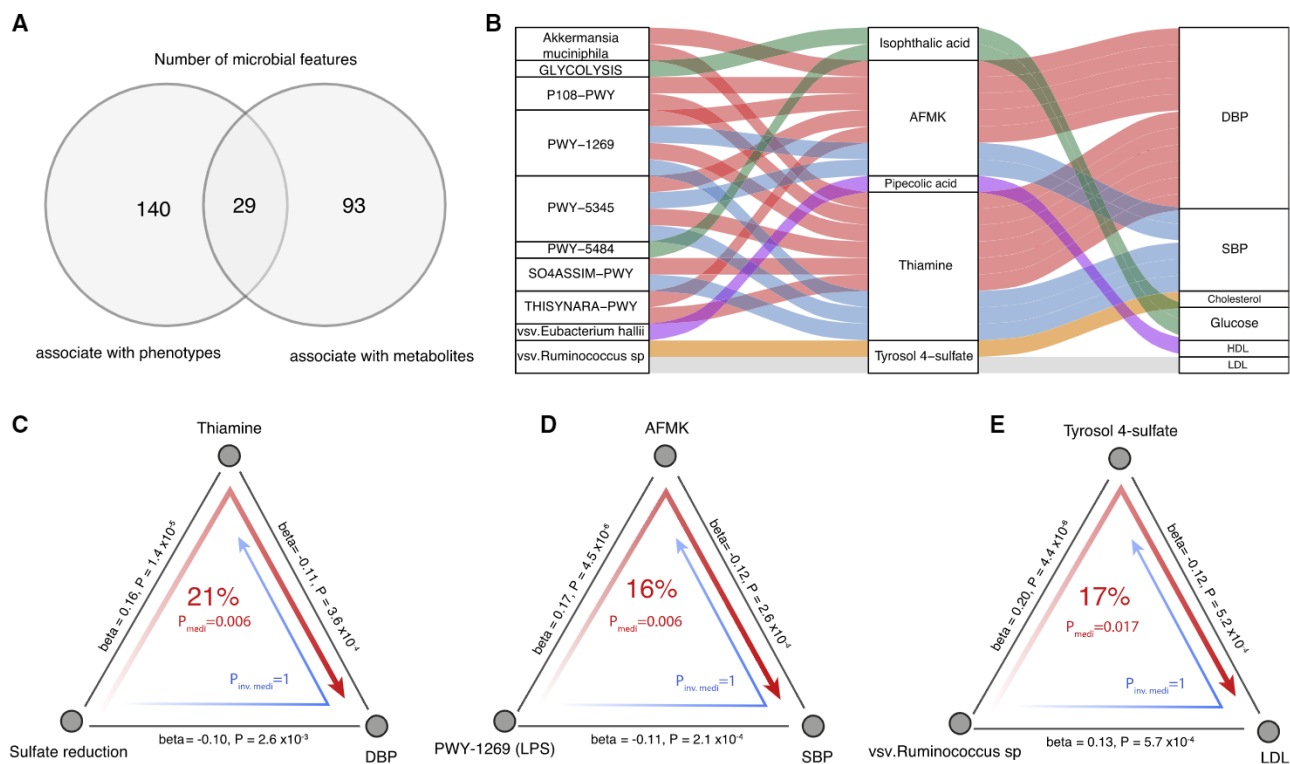


Figure 7. Mediation linkages among the gut microbiome, metabolites, and phenotypes

(A) Venn plot of the number of microbial features that were associated with human phenotypes and plasma metabolites, respectively.

(B) Parallel coordinates chart showing the 21 significant mediation effects of plasma metabolites. The left panel shows the microbial factors, the middle panel shows the plasma metabolites, and the right panel shows the phenotypes. The curved lines across panels indicate the mediation effects, while the colors correspond to different phenotypes.

(C) The microbial sulfate reduction pathway causally contributed to diastolic blood pressure through thiamine ($P_{\text{mediation}} = 0.006$, 21%).

(D) Microbial lipopolysaccharides pathway causally contributed to systolic blood pressure through AFMK ($P_{\text{mediation}} = 0.006$, 16%).

(E) A *Ruminococcus* sp. vSV (300–305 kb) causally contributed to plasma LDL through tyrosol 4-sulfate ($P_{\text{mediation}} = 0.017$, 17%). Inverse mediation was performed to check whether human phenotypes can influence the microbiome through metabolites. The gray lines indicate the delta associations, with corresponding normalized beta values and p values. The red arrowed lines indicate the microbial effects on phenotypes mediated by metabolites, with the corresponding mediation p values. The blue arrowed lines indicate reverse mediation effects, i.e., the microbial effects on metabolites mediated by phenotypes, with the corresponding inverse mediation p values.

See also Table S4.

pathway, such as cysteine, are essential for bacterial thiamine (vitamin B1) biosynthesis (Begley, 1996), and lipopolysaccharides can activate melatonin oxidation to produce AFMK (Silva et al., 2004).

We also identified several metabolite mediation effects in the microbial impact on plasma lipids and glucose levels (Figure 7B). An interesting example here is tyrosol 4-sulfate, a uremic toxin that mediates the effect of a vSV in *Ruminococcus* sp. (300–305 kb) on plasma levels of LDL (17%, $P_{\text{mediation}} = 0.017$, Figure 7E). This vSV contains an ATPase that is responsible for transmembrane transport of metabolites (Aguilar-Barajas et al., 2011).

Significant increase of microbial antibiotic resistance

The increase of antibiotic resistance can pose a great burden in fighting infectious diseases, while virulence factors are essential for the commensal microbiota to maintain their colonization niche and evade the host's immune response. We further sys-

tematically characterized and compared the abundances of 29 antibiotic resistance genes and 59 virulence genes over time. We observed a significant increase in the total antibiotic resistance gene load ($p = 1.1 \times 10^{-9}$) and a decrease in the total number of virulence genes ($p = 5.1 \times 10^{-4}$) (Figures S7A and S7B). At the individual-gene level, 55.2% (16 out of 29) of antibiotic resistance genes and 18.6% (11 out of 59) of virulence genes showed a significant difference ($FDR < 0.05$) between time points (Table S2). Specifically, 15 out of 16 antibiotic resistance genes showed an increase in their abundance, with the highest change observed for tetracycline resistance genes (Table S2) such as tetracycline resistance protein Q (TetQ), which is widely distributed in *Bacteroides* species (Veloo et al., 2019). In line with this, the increase of tetracycline resistance gene abundance was associated with the increased abundance of multiple *Bacteroides* species (e.g., *B. vulgatus*, *B. uniformis*, and *B. ovatus*, Figure S7C; Table S4) whose abundance also increased in the follow-up (Table S2).

Although antibiotic prescription in the Netherlands remains the lowest in the Europe, tetracycline, aminoglycoside, and lincosamide are among the top broad-spectrum veterinary antibiotics (Havelaar et al., 2017) and may contribute to increased microbial antibiotic resistance in humans (Aslam et al., 2018). We therefore examined the correlation of baseline meat intake with the changes in the abundance of microbial antibiotic resistance genes and found positive associations with aminoglycoside resistance ($r_{\text{Spearman}} = 0.18$, $p = 9.2 \times 10^{-4}$) and lincosamide resistance ($r_{\text{Spearman}} = 0.15$, $p = 5.5 \times 10^{-3}$) (Figures S7D and S7E; Table S4). These observations raise concerns about antibiotic usage in farming, which may be contributing to the spread of microbial antibiotic resistance in the human gut ecosystem.

DISCUSSION

Over the past years, numerous associations between a disrupted microbiota and diseases, for example, gastrointestinal and cardiometabolic diseases, have been observed in large cross-sectional studies (Chen et al., 2020a, 2020b; Falony et al., 2016; Rothschild et al., 2018; Vieira-Silva et al., 2020; Zhernakova et al., 2016). However, the key to understanding the role of a disrupted microbiota in human diseases is to answer how stable the microbiota is and whether within-individual microbial differences can be linked to changes in host health status. We therefore systematically characterized the microbial changes at both compositional and genomic level at two time points 4 years apart in 338 individuals from the Lifelines-DEEP cohort. We observed that the gut microbiome showed a degree of long-term within-individual stability in both microbial abundance and microbial genome. In particular, we found that the genetic makeup of microbes shows individuality that can be used as a fingerprint to distinguish metagenomic samples belonging to the same individual. In addition, the longitudinal association analysis between changes in gut microbiome and host phenotypes and plasma metabolites revealed *in silico* causal relationships and putative mechanistic insights into the importance of the gut microbiome in human health and disease onset (depression, IBS, and asthma). Finally, we observed a significant increase in antibiotic resistance in the gut microbiome between the two time points that cannot be explained by self-reported antibiotic intake. Our study provides evidence that antibiotic exposure via meat consumption may result in an increase in the antibiotic resistance genes in the human gut microbiome.

Previous investigations on the short-term (within 1 year) temporal stability of microbial composition and genetic makeup in adults revealed that metagenomic samples obtained from the same individual are more similar to one another than to those from different individuals (Garud et al., 2019; Mehta et al., 2018). Our large-scale characterization of the longer-term (4-year) stability of the gut microbiome extends this observation. In addition, we found that within-individual differences in gut microbial composition were smaller in participants who had a higher alpha-diversity at baseline, supporting the hypothesis that microbial communities with higher diversity tend to be more stable over time (Coyte et al., 2015).

We also observed that the genetic stability of gut microbes varies substantially across different species, and a set of species

from (but not limited to) the genera *Bacteroides*, *Bifidobacterium*, *Methanobrevibacter*, and *Phascolarctobacterium* showed relatively high within-individual stability over a long period of time. Notably, previous studies showed that some of these species, e.g., *Bacteroides* and *Bifidobacterium*, are colonized in early life (Yassour et al., 2018) and show high genetic stability in childhood (Vatanen et al., 2019). These data suggest that each person is likely to have individual-specific microbial genetic components that are distinct from those of others and may persist from childhood to adulthood. The gut microbial genetic profile can therefore serve as a host fingerprint to uniquely distinguish stool samples belonging to the same host. In this study, we constructed a microbial fingerprinting model that combines 30 microbial features, including microbial composition, SNP profiles of 13 species, and SVs of 16 species. This model showed 82% accuracy when identifying Lifelines-DEEP samples taken 4 years apart and 95% accuracy when identifying HMP samples taken 1 year apart. These results demonstrate a potential application of our method in distinguishing sample mix-ups but also raise potential privacy concerns for subjects enrolled in human microbiome research projects.

Characterization of the long-term changes in the gut microbiome is crucial for understanding the role of the gut microbiome in chronic diseases, which have long duration and generally slow progression. Differential microbial abundances have been characterized for around half of microbial species and pathways, and within-individual differences in microbial genetic makeup have also been observed. Interestingly, the bacterial SNP haplotype and SV changes in our study did not associate with abundance changes, which reveals an extra layer of information about the microbiome's contribution to host health that is independent of abundance alterations. In contrast to previous studies that only focused on associations of microbial abundances to host phenotypes (Lloyd-Price et al., 2019; Vatanen et al., 2018; Zhou et al., 2019), microbial genetic associations that connect genomic variations with genetically encoded functions to phenotypic changes can provide putative mechanistic information. We noted that increased or decreased abundances of the species that we observed to be genetically unstable (*R. torques*, *S. parasanguinis*, and *F. prausnitzii*) have also been associated with various human diseases (Joossens et al., 2011; Ray et al., 2014; Vacca, 2017; Zhernakova et al., 2016). *F. prausnitzii* has been shown to support mucosal immune homeostasis (Hornef and Pabst, 2016), mostly linked to its capacity for butyrate production (Miquel et al., 2013). However, our data show that a greater increase in a vSV of *F. prausnitzii* (1,962–1,966 kb, Table S4) was associated with a lower increase in the number of leucocyte cells, and this SV region encodes multiple riboflavin biosynthesis genes. Interestingly, we also observed *F. prausnitzii* strain replacement in 37 individuals and established many associations with plasma metabolites that affect host immunity, including with licoriso flavan A and p-cresol sulfate from the class of isoflavonoids and uremic toxins. Taken together, our data suggest mechanisms underlie the role of *F. prausnitzii* in host immunity.

Notably, metabolite associations to the gut microbiome were significantly enriched for uremic toxins and thiamine. Gut microbiota derive uremic toxins from dietary protein, and the

accumulation of uremic toxins can induce chronic sterile inflammation, which in turn increases the risks of kidney and cardiometabolic diseases (Solomon et al., 2010). We characterized 58 protein-binding uremic toxins and detected 97 microbial associations for 16 uremic toxins. One of the most-associated uremic toxins is hippuric acid, a cardiometabolic risk-related metabolite that can significantly contribute to the prediction of weight gain (Yu et al., 2018; Zhao et al., 2016). We observed several microbial associations with hippuric acid, such as associations between *B. wexlerae* SVs and hippuric acid. These *B. wexlerae* SVs were also associated with BMI, suggesting that *B. wexlerae* may contribute to metabolic disorder, potentially through hippuric acid metabolism.

Vitamin B1 production is dependent on the gut microbiome, and B1 deficiency can impact the cardiovascular system (DiNicolantonio et al., 2013). Among the microbial associations to vitamin B1, our top association was related to the abundance of *Akkermansia muciniphila*, a well-known beneficial microbe that controls gut barrier function and homeostatic functions (Everard et al., 2013). Our mediation analysis identified 21 relationships, suggesting that metabolites can mediate the microbial impact on host phenotype, particularly for cardiometabolic traits. With this analysis, we further revealed that *A. muciniphila* may influence blood pressure through vitamin B1 production, providing a rationale for the development of treatment that uses this human mucus colonizer for the prevention of hypertension. Altogether, our longitudinal microbial associations and mediation analyses on host phenotypes and plasma metabolites reveal functional insights and putative causality for the role of the gut microbiome in human health and disease onset.

Limitations of study

We acknowledge several limitations in our study. The study sampled fecal samples 4 years apart in 338 individuals, making it the longest duration metagenomic sequencing-based longitudinal microbiome study with the largest sample size to date. We systematically investigated the compositional and genetic variation over time and linked microbial changes to phenotypic changes. We observed that a gut microbial composition with higher baseline diversity is more stable over time. The same finding was observed in a recent longitudinal study that followed the gut microbiome in 1,282 individuals for 5 years using 16 s rRNA sequencing (Frost et al., 2020). Replication in independent cohorts with longer duration and larger sample size may further strengthen this observation and underline biological significance. The Lifelines-DEEP cohort comprises Dutch participants of Caucasian ethnicity from the northern region of the Netherlands. It is thus possible that Lifelines-DEEP results are biased toward a region-specific microbial background constrained by host genetics and local environmental exposures. We further validated the performance of our microbial fingerprinting model in the HMP cohort that was dominated by Caucasians living in the US. It would be interesting to see whether the model performs equally well for cohorts with more diverse or other specific ethnic backgrounds. Furthermore, longitudinal associations are not proof of causation, although we did carry out causal mediation analysis to infer *in silico* causality. We primarily focused on biologically plausible mechanisms by integrating the longitudinal

metabolism dataset in order to provide mechanistic hypotheses that pinpoint specific microbial genetics and function but also demonstrate which metabolites are likely to mediate the impact of the gut microbiome on the host phenotype. Experimental validation is thus needed.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Human subjects
- METHOD DETAILS
 - Metagenomic data generation and preprocessing
 - Taxonomic profiles
 - Functional profiles
 - Antibiotic resistance genes
 - Virulence genes
 - Strain level SNP haplotypes
 - Structural variants in microbial genome
 - Plasma untargeted metabolomics
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Principal coordinates analysis (PCoA)
 - Comparison of microbial composition dissimilarity
 - Differential microbiome feature abundance
 - Distance matrix-based individual classification
 - Stepwise distance matrices combination
 - Microbial associations to host phenotypes and metabolites
 - Mediation linkage inference

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cell.2021.03.024>.

ACKNOWLEDGMENTS

We thank the participants and staff of the Lifelines cohort for their collaboration, particularly B. Bolmer and S. Gerritsma for coordinating the Lifelines data. We thank J. Dekens and J. Arends for management and technical support and K. McIntyre for English editing. We also thank the Genomics Coordination Center for providing data infrastructure and access to high-performance computing clusters. This project was funded by the Netherlands Heart Foundation (IN-CONTROL CVON grant 2012-03 and 2018-27 to A.Z. and J.F.); the NWO Gravitation Exposome-NL (024.004.017) to J.F., A.K., and A.Z.; NWO-VIDI 864.13.013 to J.F., NWO-VIDI 016.178.056 to A.Z., NWO-VIDI 016.136.308 to R.K.W., and NWO Spinoza Prize SPI 92-266 to C.W.; the European Research Council (ERC) FP7/2007-2013/ERC Advanced Grant 2012-322698 to C.W., ERC Starting Grant 715772 to A.Z., and ERC Consolidator grant 101001678 to J.F.; and the RuG Investment Agenda Grant Personalized Health to C.W. Furthermore, J.F. and C.W. were supported by the Netherlands Organ-on-Chip Initiative, an NWO Gravitation project (024.003.001) funded by the Ministry of Education, Culture and Science of the government of the Netherlands. L.C. also holds a Foundation De

Cock-Hadders grant (20:20-13) and a joint fellowship from the University Medical Center Groningen and China Scholarship Council (CSC201708320268). D.W. holds a fellowship from the China Scholarship Council (CSC201904910478). S.G. holds a fellowship from the Graduate School of Medical Sciences, University of Groningen. T.S. holds a scholarship from the Junior Scientific Masterclass (JSM), University of Groningen. R.K.W. is supported by the Seerave Foundation and the Dutch Digestive Foundation (16-14). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

C.W., A.Z., and J.F. conceptualized and managed the study. L.C., D.W., S.G., A.K., A.V.V., R.G., and T.S. generated the data. L.C., D.W., S.G., and A.K. analyzed the data. L.C., D.W., S.G., and J.F. drafted the manuscript. L.C., D.W., S.G., A.K., A.V.V., R.G., T.S., E.S., R.K.W., C.W., A.Z., and J.F. reviewed and edited the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 9, 2020

Revised: December 2, 2020

Accepted: March 11, 2021

Published: April 9, 2021

REFERENCES

- Aguilar-Barajas, E., Díaz-Pérez, C., Ramírez-Díaz, M.I., Riveros-Rosas, H., and Cervantes, C. (2011). Bacterial transport of sulfate, molybdate, and related oxyanions. *Biometals* *24*, 687–707.
- Aitchison, J. (1982). The statistical analysis of compositional data. *J. R. Stat. Soc. B* *44*, 139–160.
- Alaei-Shahmiri, F., Soares, M.J., Zhao, Y., and Sherriff, J. (2015). The impact of thiamine supplementation on blood pressure, serum lipids and C-reactive protein in individuals with hyperglycemia: a randomised, double-blind cross-over trial. *Diabetes Metab. Syndr.* *9*, 213–217.
- Aslam, B., Wang, W., Arshad, M.I., Khurshid, M., Muzammil, S., Rasool, M.H., Nisar, M.A., Alvi, R.F., Aslam, M.A., Qamar, M.U., et al. (2018). Antibiotic resistance: a rundown of a global crisis. *Infect. Drug Resist.* *11*, 1645–1658.
- Bateman, A., Martin, M.J., O'Donovan, C., Magrane, M., Apweiler, R., Alpi, E., Antunes, R., Ar-Ganiska, J., Bely, B., Bingley, M., et al.; UniProt Consortium (2015). UniProt: a hub for protein information. *Nucleic Acids Res.* *43*, D204–D212.
- Begley, T.P. (1996). The biosynthesis and degradation of thiamin (vitamin B1). *Nat. Prod. Rep.* *13*, 177–185.
- Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., and Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.* *125*, 279–284.
- Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C.A., Keseler, I.M., Kothari, A., Krummenacker, M., Latendresse, M., Mueller, L.A., et al. (2016). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* *44* (D1), D471–D480.
- Caspi, R., Billington, R., Fulcher, C.A., Keseler, I.M., Kothari, A., Krummenacker, M., Latendresse, M., Midford, P.E., Ong, Q., Ong, W.K., et al. (2018). The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res.* *46* (D1), D633–D639.
- Chen, L., Garmeva, S., Zernakova, A., Fu, J., and Wijmenga, C. (2018). A system biology perspective on environment-host-microbe interactions. *Hum. Mol. Genet.* *27* (R2), R187–R194.
- Chen, L., Collij, V., Jaeger, M., van den Munckhof, I.C.L., Vich Vila, A., Kurilshikov, A., Gacesa, R., Sinha, T., Oosting, M., Joosten, L.A.B., et al. (2020a). Gut microbial co-abundance networks show specificity in inflammatory bowel disease and obesity. *Nat. Commun.* *11*, 4018.
- Chen, L., van den Munckhof, I.C.L., Schraa, K., Ter Horst, R., Koehorst, M., van Faassen, M., van der Ley, C., Doestzada, M., Zernakova, D.V., Kurilshikov, A., et al.; Human Functional Genomics Project (2020b). Genetic and Microbial Associations to Plasma and Fecal Bile Acids in Obesity Relate to Plasma Lipids and Liver Fat Content. *Cell Rep.* *33*, 108212.
- Coyte, K.Z., Schluter, J., and Foster, K.R. (2015). The ecology of the microbiome: Networks, competition, and stability. *Science* *350*, 663–666.
- DiNicolantonio, J.J., Niaz, A.K., Lavie, C.J., O'Keefe, J.H., and Ventura, H.O. (2013). Thiamine supplementation for the treatment of heart failure: a review of the literature. *Congest. Heart Fail.* *19*, 214–222.
- Duranton, F., Cohen, G., De Smet, R., Rodriguez, M., Jankowski, J., Vanholder, R., and Argiles, A.; European Uremic Toxin Work Group (2012). Normal and pathological concentrations of uremic toxins. *J. Am. Soc. Nephrol.* *23*, 1258–1270.
- Everard, A., Belzer, C., Geurts, L., Ouwerkerk, J.P., Druart, C., Bindels, L.B., Guiot, Y., Derrien, M., Muccioli, G.G., Delzenne, N.M., et al. (2013). Cross-talk between *Akkermansia muciniphila* and intestinal epithelium controls diet-induced obesity. *Proc. Natl. Acad. Sci. USA* *110*, 9066–9071.
- Faith, J.J., Guruge, J.L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A.L., Clemente, J.C., Knight, R., Heath, A.C., Leibel, R.L., et al. (2013). The long-term stability of the human gut microbiota. *Science* *341*, 1237439.
- Falony, G., Joossens, M., Vieira-Silva, S., Wang, J., Darzi, Y., Faust, K., Kurilshikov, A., Bonder, M.J., Valles-Colomer, M., Vandeputte, D., et al. (2016). Population-level analysis of gut microbiome variation. *Science* *352*, 560–564.
- Franzosa, E.A., Huang, K., Meadow, J.F., Gevers, D., Lemon, K.P., Bohannan, B.J., and Huttenhower, C. (2015). Identifying personal microbiomes using metagenomic codes. *Proc. Natl. Acad. Sci. USA* *112*, E2930–E2938.
- Franzosa, E.A., McIver, L.J., Rahnavard, G., Thompson, L.R., Schirmer, M., Weingart, G., Lipson, K.S., Knight, R., Caporaso, J.G., Segata, N., and Huttenhower, C. (2018). Species-level functional profiling of metagenomes and metatranscriptomes. *Nat. Methods* *15*, 962–968.
- Frost, F., Kacprowski, T., Rühlemann, M., Pietzner, M., Bang, C., Franke, A., Nauck, M., Völker, U., Völzke, H., Dörr, M., et al. (2020). Long-term instability of the intestinal microbiome is associated with metabolic liver disease, low microbiota diversity, diabetes mellitus and impaired exocrine pancreatic function. *Gut* *70*, 522–530.
- Fuhrer, T., Heer, D., Begemann, B., and Zamboni, N. (2011). High-throughput, accurate mass metabolome profiling of cellular extracts by flow injection-time-of-flight mass spectrometry. *Anal. Chem.* *83*, 7074–7080.
- Garud, N.R., Good, B.H., Hallatschek, O., and Pollard, K.S. (2019). Evolutionary dynamics of bacteria in the gut microbiome within and across hosts. *PLoS Biol.* *17*, e3000102.
- Greenblum, S., Carr, R., and Borenstein, E. (2015). Extensive strain-level copy-number variation across human gut microbiome species. *Cell* *160*, 583–594.
- Havelaar, A.H., Graveland, H., van de Kasstele, J., Zomer, T.P., Veldman, K., and Bouwknegt, M. (2017). A summary index for antimicrobial resistance in food animals in the Netherlands. *BMC Vet. Res.* *13*, 305.
- Hornef, M.W., and Pabst, O. (2016). Real friends: *Faecalibacterium prausnitzii* supports mucosal immune homeostasis. *Gut* *65*, 365–367.
- Joossens, M., Huys, G., Cnockaert, M., De Preter, V., Verbeke, K., Rutgeerts, P., Vandamme, P., and Vermeire, S. (2011). Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut* *60*, 631–637.
- Kaminski, J., Gibson, M.K., Franzosa, E.A., Segata, N., Dantas, G., and Huttenhower, C. (2015). High-Specificity Targeted Functional Profiling in Microbial Communities with ShortBRED. *PLoS Comput. Biol.* *11*, e1004557.
- Kikuchi, K., Saigusa, D., Kanemitsu, Y., Matsumoto, Y., Thanai, P., Suzuki, N., Mise, K., Yamaguchi, H., Nakamura, T., Asaji, K., et al. (2019). Gut microbiome-derived phenyl sulfate contributes to albuminuria in diabetic kidney disease. *Nat. Commun.* *10*, 1835.

- Langmead, B., Wilks, C., Antonescu, V., and Charles, R. (2019). Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* 35, 421–432.
- Liu, B., Zheng, D., Jin, Q., Chen, L., and Yang, J. (2019). VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* 47 (D1), D687–D692.
- Lloyd-Price, J., Mahurkar, A., Rahnavard, G., Crabtree, J., Orvis, J., Hall, A.B., Brady, A., Creasy, H.H., McCracken, C., Giglio, M.G., et al. (2017). Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* 550, 61–66.
- Lloyd-Price, J., Arze, C., Ananthakrishnan, A.N., Schirmer, M., Avila-Pacheco, J., Poon, T.W., Andrews, E., Ajami, N.J., Bonham, K.S., Brislawn, C.J., et al.; IBDMDB Investigators (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* 569, 655–662.
- Mayo, J.C., Sainz, R.M., Tan, D.X., Hardeland, R., Leon, J., Rodriguez, C., and Reiter, R.J. (2005). Anti-inflammatory actions of melatonin and its metabolites, N1-acetyl-N2-formyl-5-methoxykynuramine (AFMK) and N1-acetyl-5-methoxykynuramine (AMK), in macrophages. *J. Neuroimmunol.* 165, 139–149.
- Mehta, R.S., Abu-Ali, G.S., Drew, D.A., Lloyd-Price, J., Subramanian, A., Lochhead, P., Joshi, A.D., Ivey, K.L., Khalili, H., Brown, G.T., et al. (2018). Stability of the human faecal microbiome in a cohort of adult men. *Nat. Microbiol.* 3, 347–355.
- Miquel, S., Martín, R., Rossi, O., Bermúdez-Humarán, L.G., Chatel, J.M., Sokol, H., Thomas, M., Wells, J.M., and Langella, P. (2013). Faecalibacterium prausnitzii and human intestinal health. *Curr. Opin. Microbiol.* 16, 255–261.
- Mishra, A.K., Gimenez, G., Lagier, J.C., Robert, C., Raoult, D., and Fournier, P.E. (2012). Genome sequence and description of *Alistipes senegalensis* sp. nov. *Stand. Genomic Sci.* 6, 1–16.
- Munukka, E., Rintala, A., Toivonen, R., Nylund, M., Yang, B., Takanen, A., Hänninen, A., Vuopio, J., Huovinen, P., Jalkanen, S., and Pekkala, S. (2017). Faecalibacterium prausnitzii treatment improves hepatic health and reduces adipose tissue inflammation in high-fat fed mice. *ISME J.* 11, 1667–1679.
- Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 405, 299–304.
- Ray, D., Alpini, G., and Glaser, S. (2014). Probiotic Bifidobacterium species: potential beneficial effects in diarrheal disorders. Focus on “Probiotic Bifidobacterium species stimulate human SLC26A3 gene function and expression in intestinal epithelial cells”. *Am. J. Physiol. Cell Physiol.* 307, C1081–C1083.
- Rezzani, R., Porter, E., De Ciuceis, C., Bonomini, F., Rodella, L.F., Paiardi, S., Boari, G.E., Platto, C., Pili, A., Avanzi, D., et al. (2010). Effects of melatonin and Pycnogenol on small artery structure and function in spontaneously hypertensive rats. *Hypertension* 55, 1373–1380.
- Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* 16, 276–277.
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C., and Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12, 77.
- Rothschild, D., Weissbrod, O., Barkan, E., Kurilshikov, A., Korem, T., Zeevi, D., Costea, P.I., Godneva, A., Kalka, I.N., Bar, N., et al. (2018). Environment dominates over host genetics in shaping human gut microbiota. *Nature* 555, 210–215.
- Schloissnig, S., Arumugam, M., Sunagawa, S., Mitreva, M., Tap, J., Zhu, A., Waller, A., Mende, D.R., Kultima, J.R., Martin, J., et al. (2013). Genomic variation landscape of the human gut microbiome. *Nature* 493, 45–50.
- Scholtens, S., Smidt, N., Swertz, M.A., Bakker, S.J., Dotinga, A., Vonk, J.M., van Dijk, F., van Zon, S.K., Wijmenga, C., Wolfenbuttel, B.H., and Stolck, R.P. (2015). Cohort Profile: LifeLines, a three-generation cohort study and biobank. *Int. J. Epidemiol.* 44, 1172–1180.
- Silva, S.O., Rodrigues, M.R., Carvalho, S.R., Catalani, L.H., Campa, A., and Ximenes, V.F. (2004). Oxidation of melatonin and its catabolites, N1-acetyl-N2-formyl-5-methoxykynuramine and N1-acetyl-5-methoxykynuramine, by activated leukocytes. *J. Pineal Res.* 37, 171–175.
- Solomon, S.D., Uno, H., Lewis, E.F., Eckardt, K.U., Lin, J., Burdman, E.A., de Zeeuw, D., Ivanovich, P., Levey, A.S., Parfrey, P., et al.; Trial to Reduce Cardiovascular Events with Aranesp Therapy (TREAT) Investigators (2010). Erythropoietic response and outcomes in kidney disease and type 2 diabetes. *N. Engl. J. Med.* 363, 1146–1155.
- Stewart, C.J., Ajami, N.J., O’Brien, J.L., Hutchinson, D.S., Smith, D.P., Wong, M.C., Ross, M.C., Lloyd, R.E., Doddapaneni, H., Metcalf, G.A., et al. (2018). Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* 562, 583–588.
- Tigchelaar, E.F., Zernakova, A., Dekens, J.A., Hermes, G., Baranska, A., Mujagic, Z., Swertz, M.A., Muñoz, A.M., Deelen, P., Cénit, M.C., et al. (2015). Cohort profile: LifeLines DEEP, a prospective, general population cohort study in the northern Netherlands: study design and baseline characteristics. *BMJ Open* 5, e006772.
- Truong, D.T., Franzosa, E.A., Tickle, T.L., Scholz, M., Weingart, G., Pasolli, E., Tett, A., Huttenhower, C., and Segata, N. (2015). MetaPhlan2 for enhanced metagenomic taxonomic profiling. *Nat. Methods* 12, 902–903.
- Truong, D.T., Tett, A., Pasolli, E., Huttenhower, C., and Segata, N. (2017). Microbial strain-level population structure and genetic diversity from metagenomes. *Genome Res.* 27, 626–638.
- Vacca, I. (2017). Biofilms: New ways for streptococci to settle down. *Nat. Rev. Microbiol.* 15, 321.
- Vatanen, T., Franzosa, E.A., Schwager, R., Tripathi, S., Arthur, T.D., Vehik, K., Lernmark, Å., Hagopian, W.A., Rewers, M.J., She, J.X., et al. (2018). The human gut microbiome in early-onset type 1 diabetes from the TEDDY study. *Nature* 562, 589–594.
- Vatanen, T., Plichta, D.R., Somani, J., Münch, P.C., Arthur, T.D., Hall, A.B., Rudolf, S., Oakeley, E.J., Ke, X., Young, R.A., et al. (2019). Genomic variation and strain-specific functional adaptation in the human gut microbiome during early life. *Nat. Microbiol.* 4, 470–479.
- Veloo, A.C.M., Baas, W.H., Haan, F.J., Coco, J., and Rossen, J.W. (2019). Prevalence of antimicrobial resistance genes in Bacteroides spp. and Prevotella spp. Dutch clinical isolates. *Clin. Microbiol. Infect.* 25, 1156.e9, 1156.
- Vich Vila, A., Imhann, F., Collij, V., Jankipersadsing, S.A., Gurry, T., Mujagic, Z., Kurilshikov, A., Bonder, M.J., Jiang, X., Tigchelaar, E.F., et al. (2018). Gut microbiota composition and functional changes in inflammatory bowel disease and irritable bowel syndrome. *Sci. Transl. Med.* 10, eaap8914.
- Vieira-Silva, S., Falony, G., Belda, E., Nielsen, T., Aron-Wisniewsky, J., Chakaroun, R., Forslund, S.K., Assmann, K., Valles-Colomer, M., Nguyen, T.T.D., et al.; MetaCardis Consortium (2020). Statin therapy is associated with lower prevalence of gut microbiota dysbiosis. *Nature* 581, 310–315.
- Wang, Z., and Zhao, Y. (2018). Gut microbiota derived metabolites in cardiovascular health and disease. *Protein Cell* 9, 416–431.
- Wijmenga, C., and Zernakova, A. (2018). The importance of cohort studies in the post-GWAS era. *Nat. Genet.* 50, 322–328.
- Wishart, D.S., Feunang, Y.D., Marcu, A., Guo, A.C., Liang, K., Vázquez-Fresno, R., Sajed, T., Johnson, D., Li, C., Karu, N., et al. (2018). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* 46 (D1), D608–D617.
- Yassour, M., Jason, E., Hogstrom, L.J., Arthur, T.D., Tripathi, S., Siljander, H., Selvenius, J., Oikarinen, S., Hyöty, H., Virtanen, S.M., et al. (2018). Strain-Level Analysis of Mother-to-Child Bacterial Transmission during the First Few Months of Life. *Cell Host Microbe* 24, 146–154.
- Yu, T.-H., Tang, W.-H., Lu, Y.-C., Wang, C.-P., Hung, W.-C., Wu, C.-C., Tsai, I.-T., Chung, F.-M., Hough, J.-Y., Lan, W.-C., and Lee, Y.J. (2018). Association between hippuric acid and left ventricular hypertrophy in maintenance hemodialysis patients. *Clin. Chim. Acta* 484, 47–51.
- Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M., and Larsen, M.V. (2012). Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* 67, 2640–2644.
- Zeevi, D., Korem, T., Godneva, A., Bar, N., Kurilshikov, A., Lotan-Pompan, M., Weinberger, A., Fu, J., Wijmenga, C., Zernakova, A., and Segal, E. (2019). Structural variation in the gut microbiome associates with host health. *Nature* 568, 43–48.

Zhao, H., Shen, J., Djukovic, D., Daniel-MacDougall, C., Gu, H., Wu, X., and Chow, W.H. (2016). Metabolomics-identified metabolites associated with body mass index and prospective weight gain among Mexican American women. *Obes. Sci. Pract.* 2, 309–317.

Zhernakova, A., Kurilshikov, A., Bonder, M.J., Tigchelaar, E.F., Schirmer, M., Vatanen, T., Mujagic, Z., Vila, A.V., Falony, G., Vieira-Silva, S., et al.; LifeLines

cohort study (2016). Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* 352, 565–569.

Zhou, W., Sailani, M.R., Contrepois, K., Zhou, Y., Ahadi, S., Leopold, S.R., Zhang, M.J., Rao, V., Avina, M., Mishra, T., et al. (2019). Longitudinal multi-omics of host-microbe dynamics in prediabetes. *Nature* 569, 663–671.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Fecal samples	This study	https://www.Lifelines.nl
Blood samples	This study	https://www.Lifelines.nl
Critical commercial assays		
AllPrep DNA/RNA Mini Kit	QIAGEN	80204
Quant-iT PicoGreen dsDNA Assay	Life Technologies	P7589
Blood Assays	Lifelines Biobank	https://www.Lifelines.nl
Software and algorithms		
R (version 3.6.0)	R Foundation	http://www.r-project.org/
Python (version 2.7.11)	Python	https://www.python.org
KneadData (version 0.4.6.1)	The Huttenhower Lab	https://huttenhower.sph.harvard.edu/kneaddata
Bowtie2 (version 2.1.0)	(Langmead et al., 2019)	http://bowtie-bio.sourceforge.net/bowtie2
MetaPhlan2 (version 2.7.2)	(Truong et al., 2015)	https://huttenhower.sph.harvard.edu/metaphlan
HUMAnN2 (version 0.10.0)	(Franzosa et al., 2018)	https://huttenhower.sph.harvard.edu/humann
ShortBRED (version 0.9.5)	(Kaminski et al., 2015)	https://huttenhower.sph.harvard.edu/shortbred
StrainPhlan (version 1.2.0)	(Truong et al., 2017)	http://segatalab.cibio.unitn.it/tools/strainphlan
ICRA	(Zeevi et al., 2019)	https://github.com/segalab/SGVFinder
SGVFinder	(Zeevi et al., 2019)	https://github.com/segalab/SGVFinder
Deposited data		
LLD baseline metagenomics	EGA	EGAD00001001991
LLD follow-up metagenomics	EGA	EGAD00001006959
HMP raw metagenomics	HMP	https://www.hmpdacc.org
LLD metabolomics	EGA	EGAD00001006953
LLD metadata	Lifelines	https://www.Lifelines.nl

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to the Lead Contact, Jingyuan Fu (j.fu@umcg.nl).

Materials availability

Both fecal and plasma samples of Lifelines participants can be requested via Lifelines biobank (<https://www.lifelines.nl/researcher>).

Data and code availability

All relevant data supporting the key findings of this study are available within the article and its supplementary information files. The raw metagenomic sequencing data and basic phenotypes (i.e., age and gender) of the Lifelines-DEEP participants at both time points are available from the European Genome-Phenome Archive (EGA, <https://ega-archive.org>) via accession numbers EGAD00001001991 and EGAD00001006959, respectively. The metabolomics data at both time points is available from the EGA via accession number EGAD00001006953. The raw metagenomic sequencing data of the Human Microbiome Project is available via <https://www.hmpdacc.org>. Due to informed consent regulations, phenotypic data of the Lifelines-DEEP cohort are available upon request to the Lifelines (<https://www.lifelines.nl/researcher>). This includes the submission of a letter of intention to the corresponding data access committee (the Lifelines Data Access Committee for the Lifelines-DEEP data (Jackie Dekens, e-mail: j.a.m.dekens@umcg.nl)). Datasets can be made available under a data transfer agreement and the data usage access is subject to local rules and regulations.

Analysis code is available via: https://github.com/GRONINGEN-MICROBIOME-CENTRE/Groningen-Microbiome/tree/master/Projects/LLDeep_microbiome_5year_follow-up.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human subjects

The Lifelines-DEEP cohort is a sub-cohort of the Lifelines biobank (167,729 participants) (Scholtens et al., 2015) that involved 1,539 participants and is being used to assess the biomedical, socio-demographic, behavioral, physical and psychological factors that contribute to health and disease from the north of the Netherlands (Tigchelaar et al., 2015; Wijmenga and Zernakova, 2018). The study has been approved by the institutional ethics review board of the University Medical Center Groningen (ref. M12.113965). This cohort has now been followed-up, with detailed phenotypic data collected at two time points around four years apart. Of the 1,135 individuals for whom we generated metagenomics sequencing data in 2013 (Zernakova et al., 2016), follow-up stool samples were collected for 338 individuals (55.6% female and 44.4% male) at the second time point. The duration between the two time points ranged from 3.33 to 3.92 years (mean = 3.53, sd = 0.12). At baseline, the mean age of participants was 48.2 years (18–80, sd = 11.7) and their mean BMI was 25.4 (17.6–43.3, sd = 4.08). At follow-up, the mean age was 51.7 years (22–84, sd = 11.7) and the mean BMI was 25.6 (16.1–37.6, sd = 4.0). Phenotypic data assessed in the present study included anthropometric traits (e.g., age, gender, BMI, height), 9 blood cell counts, 7 plasma metabolites (e.g., glucose, cholesterol, triglycerides), 12 diseases (IBS, depression, arthrosis) and 14 medications (e.g., proton pump inhibitors, oral contraceptives, beta blockers, statins).

METHOD DETAILS

Metagenomic data generation and preprocessing

Stool sample collection and processing followed the same protocol at both time points. All participants were asked to collect fecal samples at home and place them in their home freezer (−20°C) within 15 minutes after production. Subsequently, a nurse visited the participant to pick up the fecal samples on dry ice and transfer them to the laboratory. Aliquots were then made and stored at −80°C until further processing. The same protocol for fecal DNA isolation and metagenomics sequencing was used at both time points. Fecal DNA isolation was performed using the AllPrep DNA/RNA Mini Kit (QIAGEN; cat. 80204). After DNA extraction, fecal DNA was sent to the Broad Institute of Harvard and MIT in Cambridge, Massachusetts, USA, where library preparation and whole genome shotgun sequencing were performed on the Illumina HiSeq platform. From the raw metagenomics sequencing data, low-quality reads were discarded by the sequencing facility and reads belonging to the human genome were removed by mapping the data to the human reference genome (version NCBI37) with KneadData (version 0.4.6.1) Bowtie2 (version 2.1.0) (Langmead et al., 2019). After filtering, the average read depth was 12.3 million for both baseline samples and follow-up samples. The read depths of all samples at both time points were identical (paired Wilcoxon test $p = 0.89$).

Taxonomic profiles

Microbial taxonomic profiles were generated using MetaPhlan2 (version 2.7.2) (Truong et al., 2015). MetaPhlan2 relies on nearly 1 million unique clade-specific marker genes identified from around 17,000 reference genomes (13,500 bacterial and archaeal, 3,500 viral and 110 eukaryotic), allowing unambiguous taxonomic assignments, accurate estimation of organismal relative abundance and species-level resolution for bacteria, archaea, eukaryotes and viruses. Microbial species present in more than 10% of the samples were included for further analyses. This yielded a list of 157 species that account for 97.81% of the taxonomic composition.

Functional profiles

Microbial functional profiles were determined using HUMAnN2 (version 0.10.0) (Franzosa et al., 2018), which maps DNA/RNA reads to a customized database of functionally annotated pan-genomes. HUMAnN2 reported the abundances of gene families from the UniProt Reference Clusters (Bateman et al., 2015) (UniRef90), which were further mapped to microbial pathways from the MetaCyc metabolic pathway database (Caspi et al., 2016; Caspi et al., 2018). Based on MetaPhlan2, HUMAnN2 can further characterize community functional profiles stratified by known (species-level) and unclassified organisms. In total, 343 microbial pathways present in more than 10% of the samples were kept for subsequent analysis, accounting for 99.98% of microbial functional composition.

Antibiotic resistance genes

Quantification of antibiotic resistance genes in metagenomics was performed using shortBRED (version 0.9.5) (Kaminski et al., 2015) with markers generated from the ResFinder database, which reports more than 1,800 different antimicrobial resistance genes (November 2018 version) (Zankari et al., 2012). In brief, ShortBRED is a pipeline to take a set of protein sequences from a target database (i.e., ResFinder), cluster them into families, build consensus sequences to represent the families, and then reduce these consensus sequences to a set of unique identifying strings (markers). The pipeline then searches for these markers in metagenomic data and determines the presence and abundance of the protein families of interest. We classified the abundance of 29 antibiotic resistance genes that were present in at least 10% of the samples.

Virulence genes

We also searched the metagenomic data for bacterial virulence genes using shortBRED (version 0.9.5) (Kaminski et al., 2015) and markers generated from virulence factors of pathogenic bacteria database (VFDB, core dataset of DNA sequences, version: November, 2018) (Liu et al., 2019). Here, we classified the abundance of 59 virulence genes that are present in at least 10% of the samples.

Strain level SNP haplotypes

Strain SNP haplotypes were generated using StrainPhlAn1 (version 1.2.0) (Truong et al., 2017). This method is based on reconstructing consensus sequence variants within species-specific marker genes and using them to estimate strain-level phylogenies. Reconstructed markers with a percentage of ambiguous bases > 20% are discarded. Consensus sequences are then trimmed by removing the first and last 50 bases because the terminal positions have lower coverage due to the limitations in mapping reads against truncated sequences (Truong et al., 2017). Next, clades with a percentage of markers < 50% are removed, and if the percentage of samples in which a marker is present is < 50%, that marker is also removed. Samples with full sequences concatenated from all markers and a percentage of gaps > 50% are removed from the alignment. Finally, we used the multiple sequence alignment file to generate a phylogenetic distance matrix that contains the pairwise nucleotide substitution rate between strains by applying the Kimura 2-parameter method from the EMBOSS package (Rice et al., 2000). Using this method, we classified the within-individual SNP haplotype difference of the dominant strain in 37 species that present in at least 5 sample pairs, and 18 of these were obtained in at least 10% of sample pairs. To identify distinct strain clusters within species, the SNP haplotype distance matrix was normalized by dividing maximal distance and hierarchical clustering was based on the complete method from R basic function *hcluster()*. Strain clusters were defined at a tree height of 0.7 (70% dissimilarity).

Structural variants in microbial genome

We applied the SGV-Finder pipeline (Zeevi et al., 2019) to classify SVs that are either completely absent in the microbial genome of some samples (deletion SVs, dSVs) or those whose coverage is highly variable across samples (variable SVs, vSVs). Prior to SV classification, an 'iterative coverage-based read assignment' algorithm was applied that resolves ambiguous read assignments to regions that are similar between different bacteria, using information on bacterial relative abundances in the microbiome, their genomic sequencing coverage and sequencing and alignment qualities (Zeevi et al., 2019). In total, we classified 6,130 SVs, including 4,333 dSVs and 1,797 vSVs from 41 microbial species that are present in at least 5 sample pairs. The SVs of 26 species could be obtained in at least 10% of sample pairs. We further calculated Canberra distance between individuals based on the dSVs and vSVs of each microbial species, respectively.

Plasma untargeted metabolomics

Plasma samples were collected and frozen at -80°C with EDTA. During extraction, plasma samples were thawed on ice, vortexed and spun down. 20 μL of plasma was combined with 180 μL of 80% methanol and vortexed for 15 s. Samples were incubated at 4°C for one hour to precipitate proteins and then spun for 30 minutes at 3,200 RCF. 100 μL of supernatant was removed and used for Flow-Injection Time-of-Flight Mass Spectrometry (FIA-TOF) analysis at General Metabolics, Inc., Boston, USA, using protocols described previously (Fuhrer et al., 2011). In total, 1,183 metabolites with annotations were involved in the analysis. The annotated metabolites cover 18 chemical categories based on the Human Metabolome Database (Table S2) (Wishart et al., 2018). The characterization of plasma protein-bound uremic toxins, including indoxyl sulfate, p-cresyl sulfate, phenyl sulfate, phenylacetic acid and hippuric acid was based on (Wang and Zhao, 2018).

QUANTIFICATION AND STATISTICAL ANALYSIS

Principal coordinates analysis (PCoA)

The relative abundances of all microbial species and pathways were included in PCoA. We applied the *vegdist()* function from the *vegan* (version 2.5.5) R package to calculate the Bray-Curtis dissimilarity matrix based on the composition at the species level. Subsequently, classical metric multidimensional scaling was carried out based on the Bray-Curtis distance matrix to obtain different principal coordinates.

Comparison of microbial composition dissimilarity

To compare the differences in overall microbial species and pathway compositions between- and within-individuals, we applied a Wilcoxon test on Bray-Curtis dissimilarity. Since the number of dissimilarities between- and within-individuals was unbalanced, we calculated an empiric P value by permuting samples of microbial species and pathway relative abundance tables for 10,000 times.

Differential microbiome feature abundance

We applied different transformation/normalization methods for the different microbial abundance datasets, depending on the type of dataset. Centered log-ratio transformation was applied to compositional data, such as the relative abundances (sum up to 1) of microbial species and functional pathways (Aitchison, 1982), and log transformation (with a pseudo count of 1 for zero values) was

applied to read count data, such as the read counts of microbial antibiotic resistance genes and virulence gene (reads per kilobase of transcript per million reads mapped, RPKMs). Within-individual differences in microbial abundance were then assessed by using paired Wilcoxon tests. The FDR was calculated with 1,000 times permutation.

Distance matrix–based individual classification

We evaluated if microbial abundance and genome information can be used for individual classification (i.e., to identify if two samples belong to the same individual). To do so, we generated Bray-Curtis distances based on microbial species and pathway relative abundance, Kimura distance based on SNP haplotype profile and Canberra distance based on SV profiles. The samples were clustered using single-linkage clustering, also known as nearest neighbor clustering. If two samples, and only those two samples, from the same individual were clustered together as the closest neighbor, we considered that they were classified correctly. We then defined the accuracy by calculating the proportion of the total number of correctly classified pairs. Finally, by establishing a specific cutoff, we could determine whether a pair of samples came from the same individual by their dissimilarity, and the cutoff affects the performance of classifier. A receiver operating characteristic curve (ROC) was drawn based on dissimilarity to reflect the specificity and sensitivity of classification using *roc()* function from *pROC* (version 1.16.1) (Robin et al., 2011).

Stepwise distance matrices combination

A total of 71 distance matrices were present in more than 10% of sample pairs, including 69 genetic distance matrices (SNP haplotype distance matrices for 18 species, dSV and vSV distance matrices for 26 species) and 2 compositional distance matrices generated by microbial species and pathways abundance. We aimed to see whether we can utilize these genetic and microbial distance matrices to classify different samples from the same individuals. Each of these distance matrices was considered as one classifier. We carried out a stepwise forward selection approach to combine multiple microbial genetic and compositional distance matrices. The cohort was randomly divided into a discovery set made up of 60% of sample pairs and a validation set with the remaining 40% of pairs. In order to combine multiple distance matrices, we first standardized and scaled all distance matrices between 0 and 1 by dividing each matrix by its largest value. In the discovery set, we assessed the accuracy of each distance matrix in classifying samples, as described above. We started with the distance matrix that had the highest accuracy, i.e., the 1st classifier. We then moved on to the model with two distance matrices by adding another distance matrix and taking the mean value of two matrices. We tested all possible combinations and chose the combination with the highest accuracy. The classifier included at the second step was considered as the 2nd classifier. This step was then repeated to include the 3rd classifier and continued until all the distance matrices were included. In this way, we generated a series of models that included different numbers of distance matrices and tested their performance in the validation set. The whole procedure of dataset splitting and feature combination was repeated 10 times, and we determined the optimal feature number *N* at which the performance no longer improved as more matrices were added. The distance-based features were prioritized by their median ranks across 10-times feature selections, then top-*N* distance matrices were selected as the optimal combination for the final classifier.

Microbial associations to host phenotypes and metabolites

We first established microbial associations to host phenotypes and metabolites (Table S1) using linear and logistic mixed-effects model (joint associations): dependent variable \sim (intercept) + independent variable + age + sex + (1| time point) + (1| subject), for continuous and binary microbial traits, respectively. We further validated these joint associations by linking microbial changes to host phenotypic and metabolic changes with a regression model (delta associations): dependent variable changes \sim (intercept) + independent variable changes + age + sex, for continuous and binary microbial traits (dSVs), respectively. For the delta association of disease and medication phenotypes, we focused on disease onsets and new drug users and selected the healthy individuals and non-drug users at both times as controls. The Benjamini-Hochberg procedure was applied to control FDR (Benjamini et al., 2001).

Mediation linkage inference

For phenotypic and metabolic associations to the same microbial feature, we first checked whether the human phenotype was associated with the metabolite using both joint and delta association models, as described above. Next, bi-directional mediation analysis was carried out using the *mediate* function from the R package mediation (version 4.5.0) to infer the causal role of the microbiome in contributing to the human phenotype through metabolites. The Benjamini-Hochberg procedure was applied to control FDR.

Supplemental figures

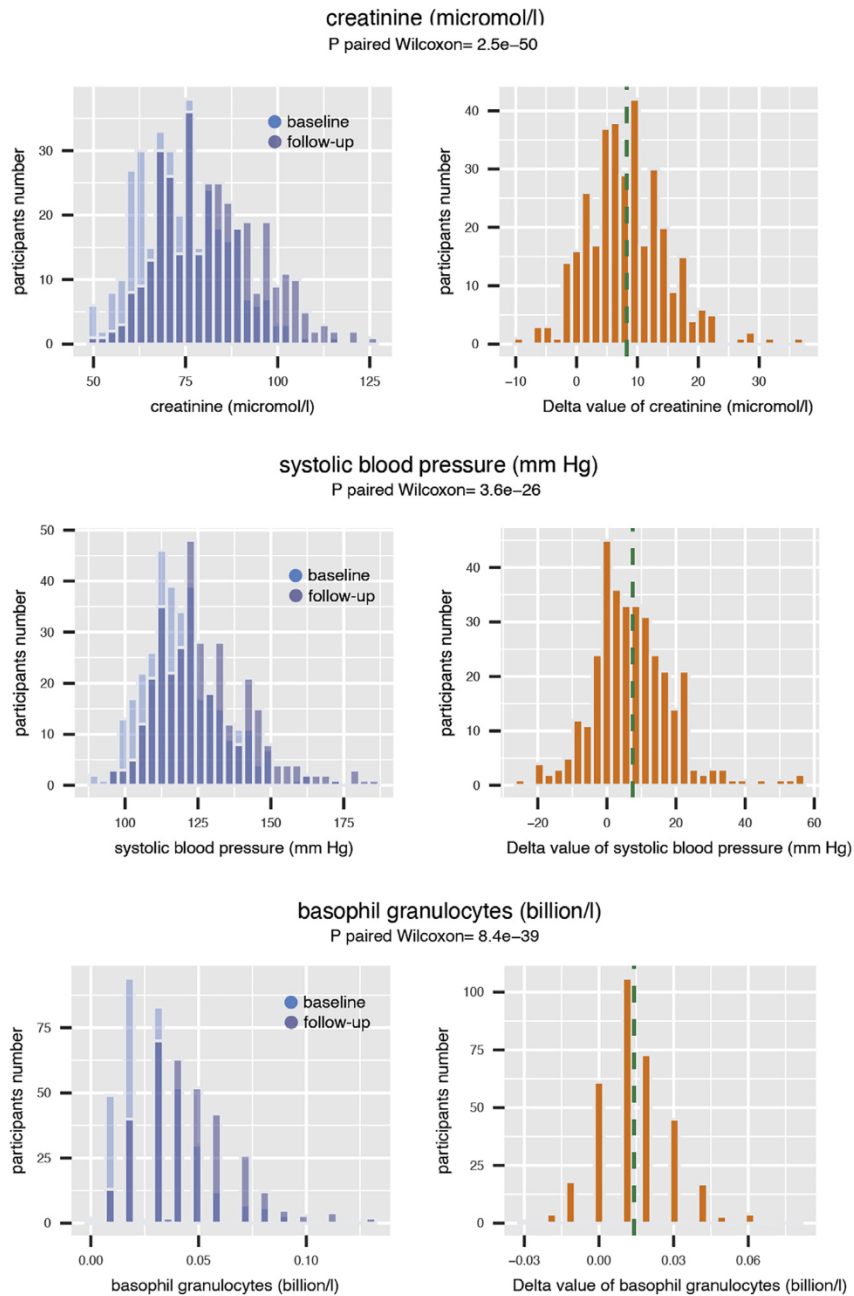


Figure S1. The distribution of phenotypes, related to Figure 4

Here we show the distributions of plasma creatinine level, systolic blood pressure and basophil granulocytes at baseline and follow-up (left panels) and the distribution of their changes (right panels). The baseline distribution is in light blue and the follow-up distribution is in dark blue. The phenotypic changes are computed as the difference between baseline and follow-up (follow-up value minus baseline value). The distribution of phenotypic changes is shown in orange, with a dashed line indicating the mean value of delta. The P value refers to the statistical significance assessed by paired Wilcoxon-test.

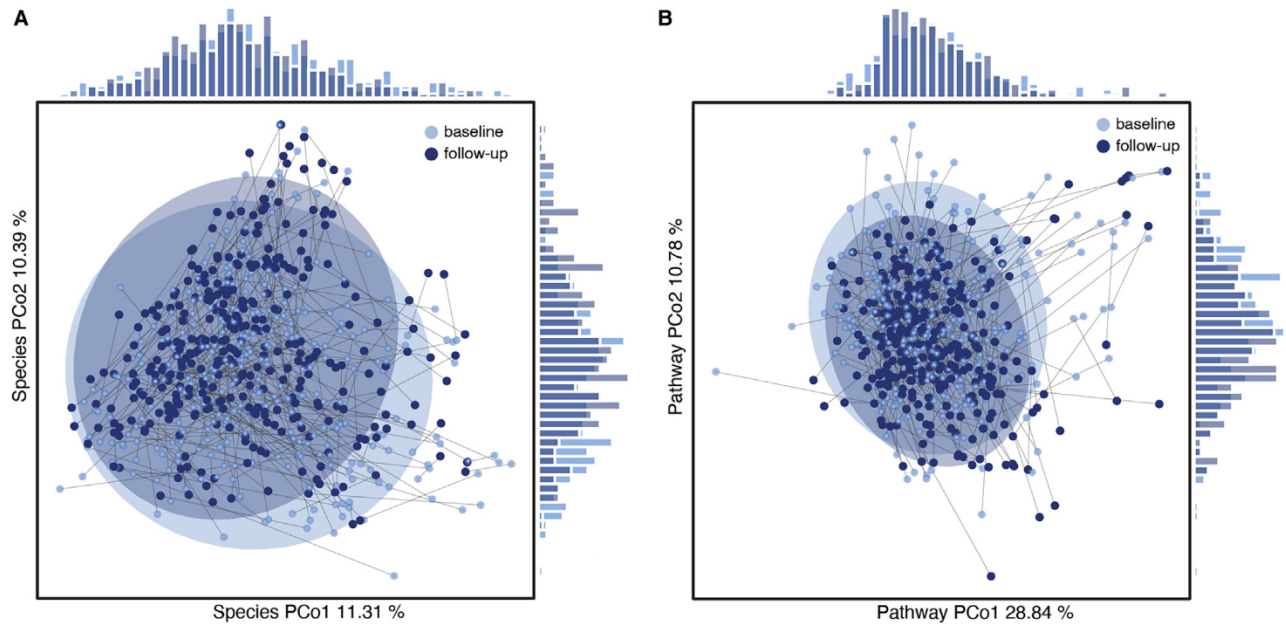


Figure S2. Comparison of microbial species and pathway compositions, related to Figure 1

Principal coordinates analysis of microbial species (A) and pathway (B) abundance (Bray-Curtis dissimilarity). The light blue dots indicate baseline samples. The dark blue dots indicate follow-up samples. Each gray line connects two samples from the same individual taken four years apart.

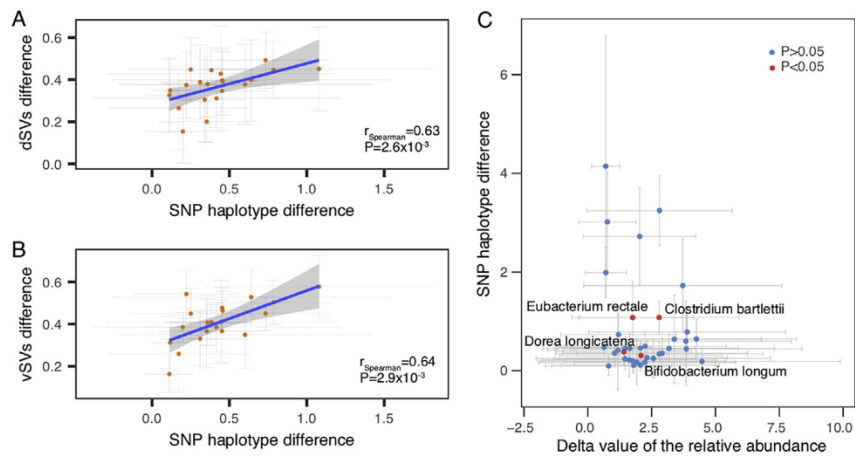


Figure S3. Comparison of microbial changes in different layers, related to Figure 2

A-B. Comparison of within-individual microbial differences in species SNPs and SVs in samples taken four years apart. Each orange dot represents one species. Cross bars represent standard deviations of within-individual SNP haplotype and SV differences. Spearman correlation was applied to assess the correlation between SNP haplotype and SV differences across species. C. Associations between the value of microbial species abundance changes and the temporal differences in SNP haplotype. Each dot represents one species. Cross bars represent the standard deviations of abundance changes (delta abundance) and SNP differences, respectively. Spearman correlation was applied to assess the correlation between the delta abundance and SNP difference for each species. Out of 37 species, only 4 species showed significant correlation at $p < 0.05$. These are highlighted as red dots and labeled with the species names.

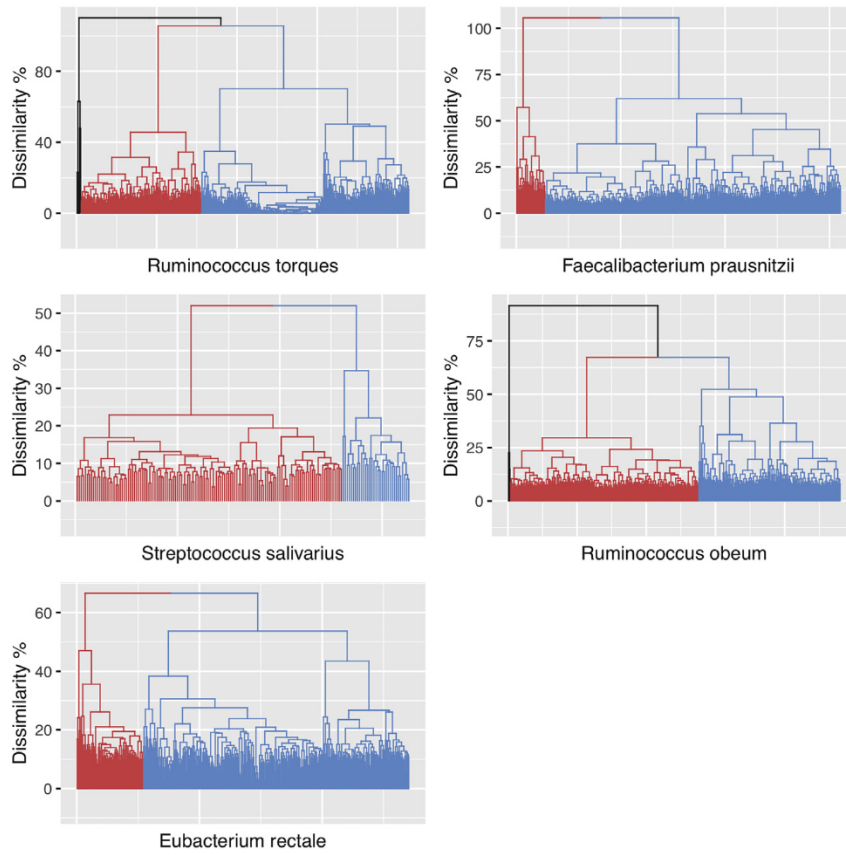
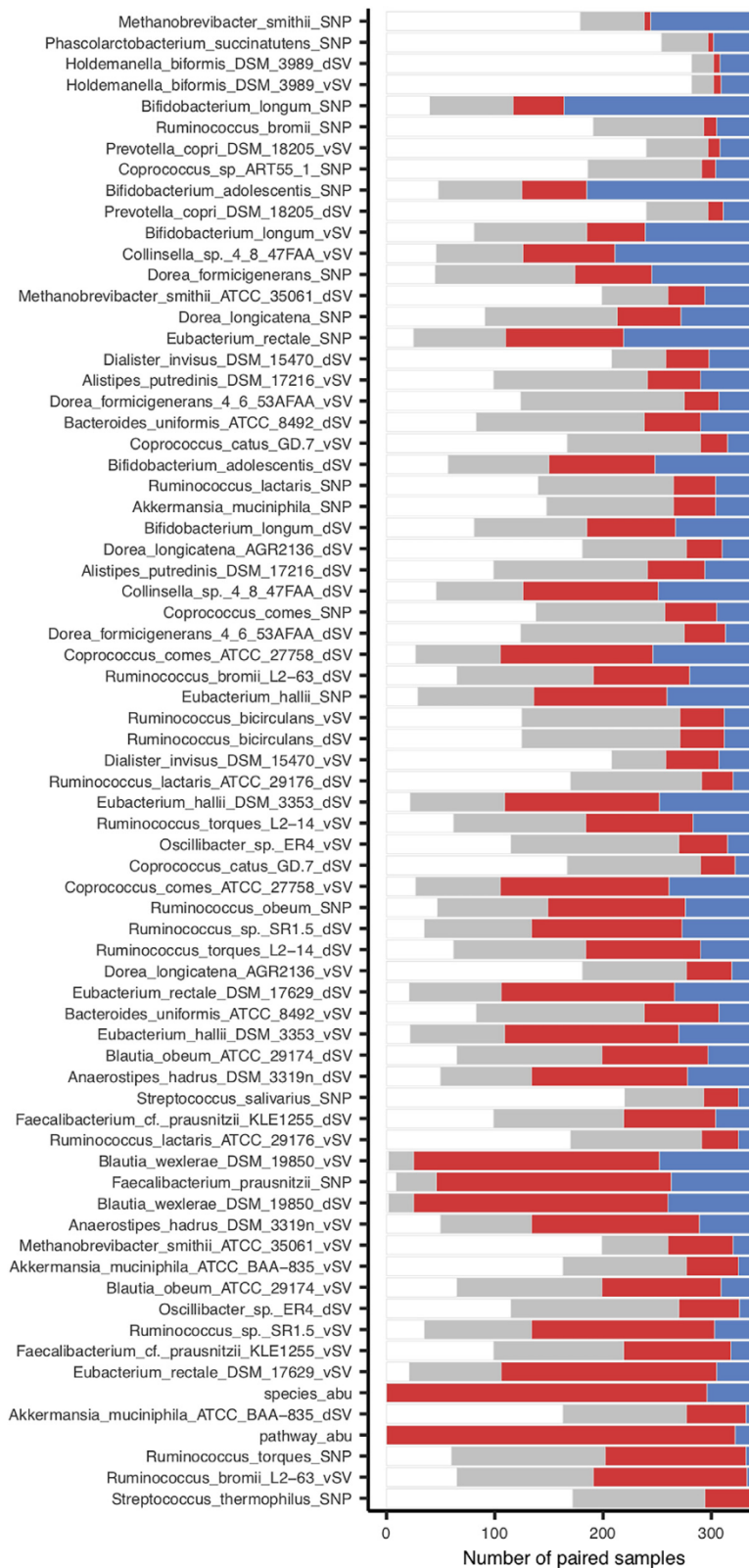


Figure S4. Distinct strain clusters based on SNP haplotype differences, related to Figure 5
 Different colors represent distinct strain clusters based on a dissimilarity of 70%.



(legend on next page)

Figure S5. Performance of microbial profiles in distinguishing samples from the same individual, related to Figure 3

Blue indicates the number of samples in which the species can be detected and the individual correctly classified at both time points. Red indicates number of samples in which the species can be detected at both time points but the individual was not correctly classified. Grey represents the number of samples in which the species can only be detected at one of the two time points. White represents the number of samples in which the species cannot be detected at either time point.

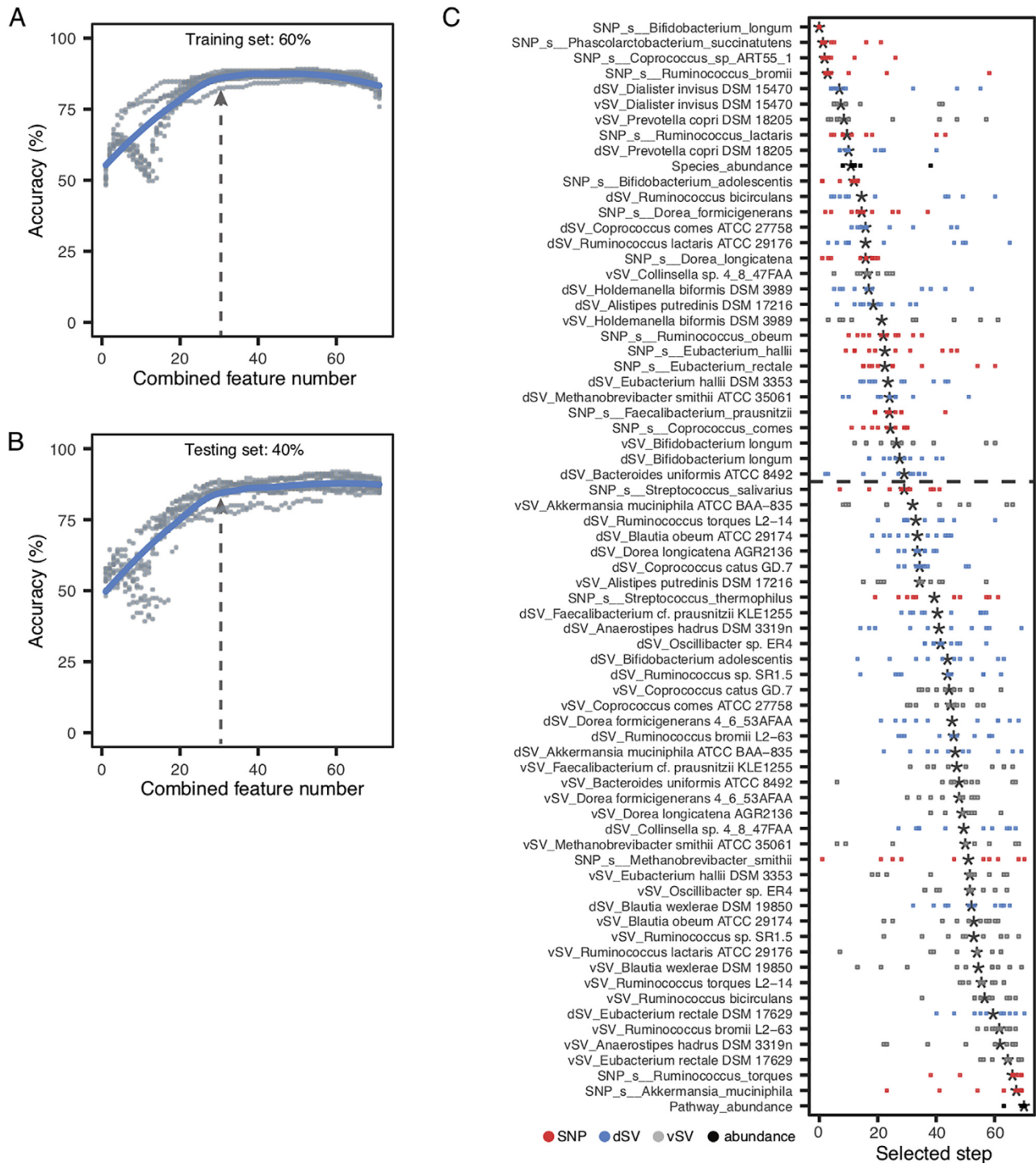


Figure S6. Performance of microbial information in classifying individuals, related to Figure 3

A. Classification accuracy in a training set made up of 60% of the metagenomic samples randomly sampled 10 times. B. Classification accuracy in the testing set containing the remaining 40% of the metagenomic samples, also randomly sampled. C. The prioritized order (selected step) of bacterial genetic and compositional profiles used in combination for microbiome fingerprinting. Dots represent the prioritized order of microbial profiles from 10-times resampling. Stars represent the median value of dots.

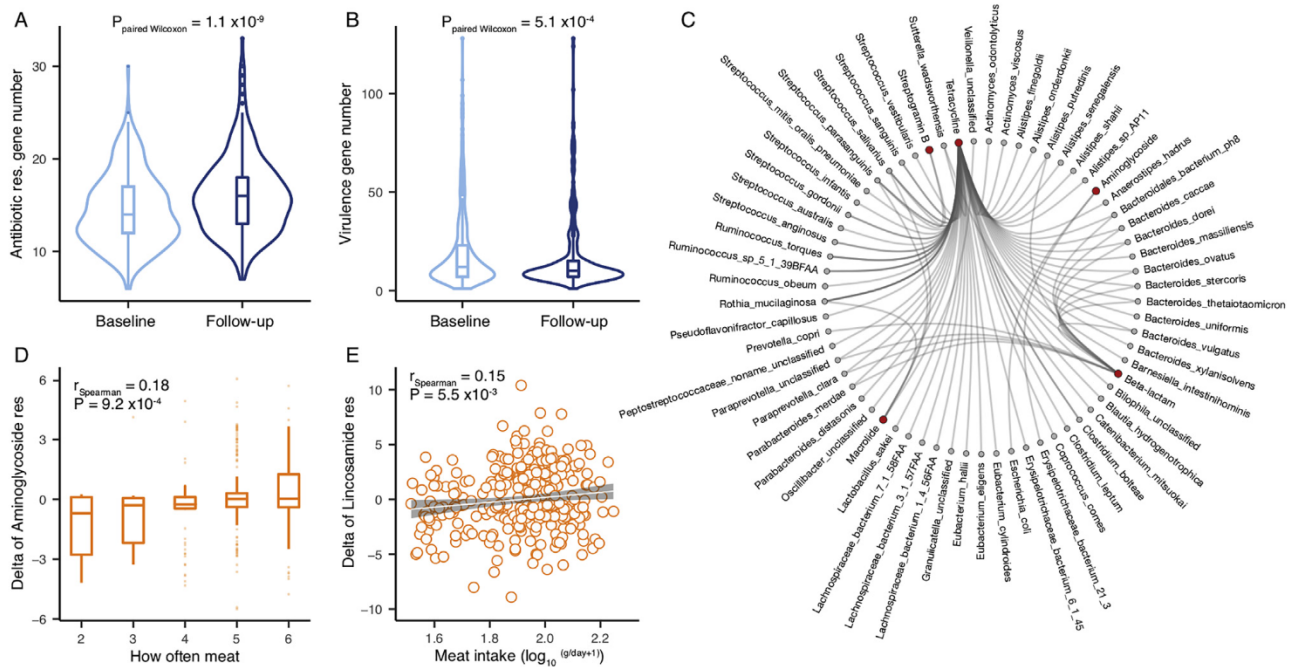


Figure S7. Long-term changes in antibiotic resistance genes and virulence factors, related to Table S2

A. Comparison of the total load of microbial antibiotic resistance genes at baseline (light blue) and follow up (dark blue). B. Comparison of the number of microbial virulence genes at baseline (light blue) and follow up (dark blue). P values from the pairwise Wilcoxon test are shown in (A) and B. C. Positive delta associations between the abundance of microbial species and the abundance of antibiotic resistance genes. Red dots indicate the corresponding antibiotics of antibiotic resistance genes. Gray dots indicate microbial species. Each line represent a significant association. D. Association of the microbial Aminoglycoside resistance gene with meat consumption. x axis shows the frequency of meat consumption. y axis shows the delta value of aminoglycoside resistance gene between two time points. The corresponding correlation coefficient and P value from the Spearman correlation are shown. E. Association of microbial Lincosamide resistance gene with meat consumption. x axis indicates meat intake. y axis indicates the delta value of the Lincosamide resistance gene between two time points. The corresponding correlation coefficient and P value from the Spearman correlation are shown.