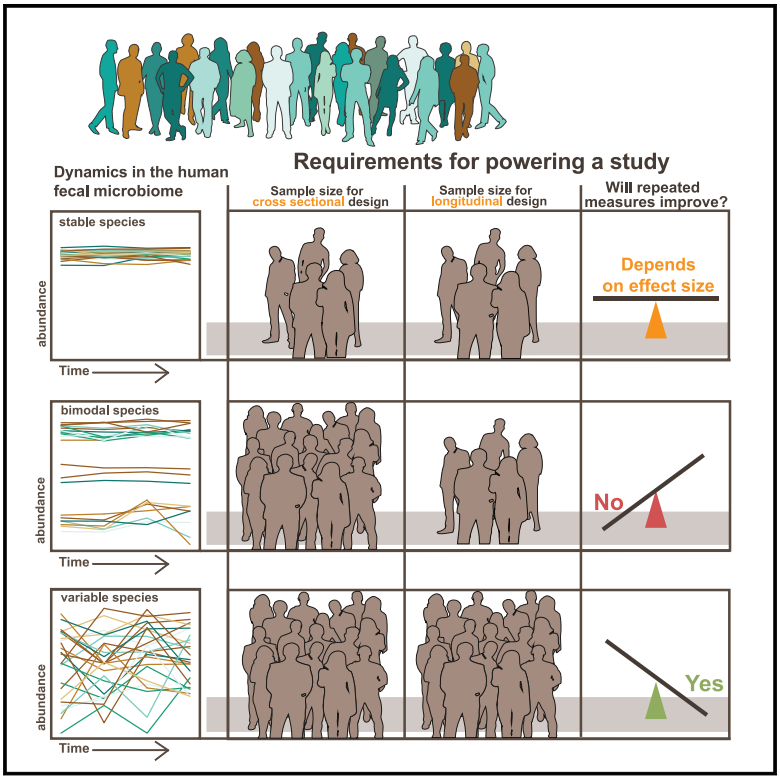
Dynamics of the normal gut microbiota: A longitudinal one-year population study in Sweden

Article

Cell Host & Microbe

Dynamics of the normal gut microbiota: A longitudinal one-year population study in Sweden

Graphical abstract



Authors

Lisa M Olsson, Fredrik Boulund, Staffan Nilsson, ..., Göran Bergström, Valentina Tremaroli, Fredrik Bäckhed

Correspondence

valentina.tremaroli@wlab.gu.se (V.T.), fredrik@wlab.gu.se (F.B.)

In brief

The intra-individual variation of the gut microbiota and its diagnostic and clinical relevance remain underappreciated. In a disease-free population, using metagenomics and 16S rRNA gene sequencing, Olsson et al. show that both composition and functional potential vary extensively within individuals. They also show that different species have different variability patterns.

Highlights

- 23% of the total compositional variation in the gut microbiome is intra-individual
- Stability markers might be specific to age and geography
- Several disease-associated species have large normal intra-individual variation
- Repeated measurements are required for reliable assessment of variable features

Olsson et al., 2022, Cell Host & Microbe *30*, 726–739
May 11, 2022
https://doi.org/10.1016/j.chom.2022.03.002

CellPress

CellPress
OPEN ACCESS

Cell Host & Microbe

Article

Dynamics of the normal gut microbiota: A longitudinal one-year population study in Sweden

Lisa M Olsson,[1] Fredrik Boulund,[2] Staffan Nilsson,[3,10] Muhammad Tanweer Khan,[1] Anders Gummesson,[1,4] Linn Fagerberg,[5] Lars Engstrand,[2,6] Rosie Perkins,[1] Mathias Uhlén,[5,7] Göran Bergström,[1,8] Valentina Tremaroli,[1,11,*] and Fredrik Bäckhed[1,8,9,11,12,*]

[1]The Wallenberg Laboratory, Department of Molecular and Clinical Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
[2]Department of Microbiology, Tumour and Cell Biology, Karolinska Institutet, Stockholm, Sweden
[3]Department of Mathematical Sciences, Chalmers University of Technology, Gothenburg, Sweden
[4]Department of Clinical Genetics and Genomics, Region Västra Götaland, Sahlgrenska University Hospital, Gothenburg, Sweden
[5]Department of Proteomics, KTH-Royal Institute of Technology, Stockholm, Sweden
[6]Clinical Genomics Facility, Science for Life Laboratory, Solna, Sweden
[7]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark
[8]Department of Clinical Physiology, Region Västra Götaland, Sahlgrenska University Hospital, Gothenburg, Sweden
[9]Novo Nordisk Foundation Center for Basic Metabolic Research, Section for Metabolic Receptology and Enteroendocrinology, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark
[10]Department of Laboratory Medicine, Institute of Biomedicine, University of Gothenburg, Gothenburg, Sweden
[11]These authors contributed equally
[12]Lead contact
*Correspondence: valentina.tremaroli@wlab.gu.se (V.T.), fredrik@wlab.gu.se (F.B.)
https://doi.org/10.1016/j.chom.2022.03.002

SUMMARY

Temporal dynamics of the gut microbiota potentially limit the identification of microbial features associated with health status. Here, we used whole-genome metagenomic and 16S rRNA gene sequencing to characterize the intra- and inter-individual variations of gut microbiota composition and functional potential of a disease-free Swedish population (n = 75) over one year. We found that 23% of the total compositional variance was explained by intra-individual variation. The degree of intra-individual compositional variability was negatively associated with the abundance of *Faecalibacterium prausnitzii* (a butyrate producer) and two *Bifidobacterium* species. By contrast, the abundance of facultative anaerobes and aerotolerant bacteria such as *Escherichia coli* and *Lactobacillus acidophilus* varied extensively, independent of compositional stability. The contribution of intra-individual variance to the total variance was greater for functional pathways than for microbial species. Thus, reliable quantification of microbial features requires repeated samples to address the issue of intra-individual variations of the gut microbiota.

INTRODUCTION

The human gut hosts a dynamic microbial ecosystem: the gut microbiota, which has an extensive metabolic repertoire and is increasingly recognized to influence host metabolism and overall physiology (Koh et al., 2016; Valdes et al., 2018). Altered gut microbiota states, generally referred to as dysbiosis, have been observed for a number of human diseases (Lynch and Pedersen, 2016). Recent advances in understanding of the ecology of the gut microbiota in health and disease have led to the concept that the microbiota could provide biomarkers for disease stratification and development of novel therapeutic strategies (Xavier, 2016).

The lack of consistent and/or specific signatures of gut microbiota dysbiosis limits our understanding of the role of gut microbiota in human diseases (Duvallet et al., 2017; Sze and Schloss, 2016). Discrepancies between studies could be explained not only by technical bias in analytical methods (Costea et al., 2017; Sinha et al., 2017) but also by stochasticity of gut microbiota responses to an altered intestinal environment (Zaneveld et al., 2017) and intrinsic properties of the gut microbiota, such as heterogeneity of composition between individuals and temporal dynamics (Poyet et al., 2019; Schmidt et al., 2018). Indeed, although large metagenomic studies revealed the presence of core taxa in normal populations from Europe, the US, and China (Falony et al., 2016; Human Microbiome Project, 2012; Li et al., 2014; Lloyd-Price et al., 2017; Qin et al., 2010) and indicated features potentially linked to human health, such as community stability (Chen et al., 2021; Faith et al., 2013) and gene richness (Le Chatelier et al., 2013), others have shown that the human gut microbiota is redundant and dynamic, fluctuating in response to external perturbations (David et al., 2014; Jackson et al., 2016; Mardinoglu et al., 2018; Salonen et al., 2014; Wu et al., 2017) and internal biological processes (Thaiss et al., 2014).

Little is known about the temporal dynamics of the gut microbiota under ecologically stable conditions (Bäckhed et al., 2012;

CellPress
OPEN ACCESS

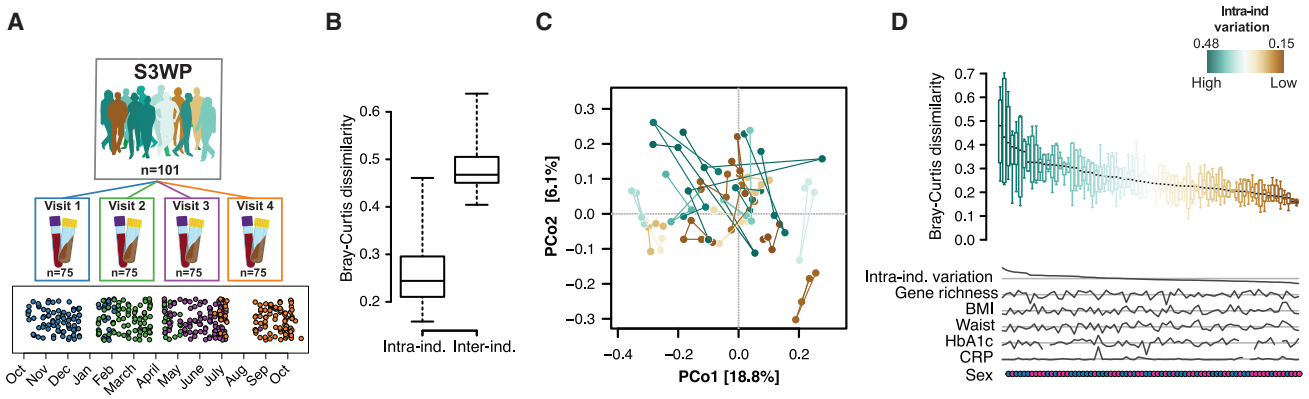


Figure 1. Compositional variability of the gut microbiota

(A) Fecal samples were obtained from 75 of the 101 individuals in the Swedish SciLifeLab SCAPIS Wellness Profiling (S3WP) study at four visits over 1 year.

(B) Boxplots of intra- and inter-individual Bray-Curtis dissimilarity for the 75 individuals (Mann-Whitney U test, $p < 2.2e{-}16$).

(C) Components 1 and 2 for dimension reduction of composition in the principal coordinates (PCo) analysis using Bray-Curtis dissimilarity for fecal species profiles. The figure shows patterns for 20 randomly selected individuals; samples from each individual are connected with lines.

(D) Compositional variability of the gut microbiota for the 75 individuals estimated by the Bray-Curtis dissimilarity. Boxplots are ordered by the median Bray-Curtis dissimilarity calculated for pairwise comparisons of the four samples from each individual (n = 6 values in each boxplot), with the brown to green scale indicating low to high gut microbiota compositional variability. Variations of gene richness, body mass index (BMI), waist, glycated hemoglobin (HbA1c), and C-reactive protein (CRP) over the 1-year study period are also indicated. Values for each variable are scaled around the mean for all the individuals. Circles indicate sex, with blue and pink representing male and female, respectively (logistic regression to median Bray-Curtis dissimilarity; $p = 0.1$). Boxes in (B) and (D) show median and interquartile ranges (IQRs); whiskers represent ±1.5 * IQR from each box's quartiles.

McBurney et al., 2019), and definitions of the normal intra- and inter-individual variations are lacking. Therefore, in this study, we characterized the intra- and inter-individual variations of the normal gut microbiota, the abundance ranges of prevalent taxa, and the variation of the functional potential in an adult Swedish population over the course of one year. We validated our findings in an independent dataset from the extended National Institutes of Health Human Microbiome Project (HMP1-II) (Lloyd-Price et al., 2017).

RESULTS

Compositional variability of the gut microbiota

To study the temporal dynamics of the normal human gut microbiota and determine the extent of intra- and inter-individual variability, we profiled the fecal microbiota, by both whole-genome metagenomics and 16S rRNA gene sequencing, of 101 men and women aged 50–65 years in the Swedish SciLifeLab SCAPIS Wellness Profiling (S3WP) study. These individuals were recruited from the Swedish CArdioPulmonary bioImage Study (SCAPIS) (Bergström et al., 2015) and screened at six visits over 2 years (Tebani et al., 2020). In the current analyses, we included the 75 individuals who did not receive antibiotics and for whom we had fecal metagenomic samples from all four visits in the first year (Figure 1A; biometric characteristics are in Table S1A).

For whole-genome metagenomics, we mapped high-quality microbial reads to a non-redundant genome catalog using the MEDUSA platform (Karlsson et al., 2014). In the downstream analyses, we included 384 abundant and prevalent microbial species that had an abundance of more than 10 counts in at least 30 of the 300 samples. We observed 184 core species that were present in all 300 samples and 316 core species when considering taxa present in at least one sample from every individual (Table S1B). Our results extend previous definitions of core species observed in cross-sectional studies (Falony et al., 2016; Qin et al., 2010).

By analyzing overall gut microbiota composition using Bray-Curtis dissimilarity, we observed less intra-individual variation than inter-individual variation (Figures 1B and 1C), confirming that the gut microbiota is individualized (Chen et al., 2021; Costello et al., 2009; Flores et al., 2014; Human Microbiome Project, 2012; Turnbaugh et al., 2009). However, 23% of the total compositional variance was explained by the intra-individual variation of the gut microbiota (permutational MANOVA, $p = 0.0001$). We did not observe any significant effect of season or visit number on the compositional variability of the gut microbiota in our population (Table S1C).

Potential stabilizing factors of the gut microbiota

We observed that the degree of compositional variability over the one-year study period differed between the individuals (Figure 1D) in line with previous results (Flores et al., 2014); individuals with the highest degree of variability had a median Bray-Curtis intra-individual dissimilarity similar to the median inter-individual dissimilarity shown in Figure 1B. We show that the degree of intra-individual compositional variability did not associate with measured anthropometric variables (e.g., sex, body mass index [BMI], glycated hemoglobin [HbA1c], and C-reactive protein [CRP]) or the macronutrient composition of the diet (Figure 1D; Tables S1D and S1E), possibly due to the homogeneous population and absence of disease. However, the intra-individual compositional variability increased significantly with decreasing relative abundance of *Faecalibacterium prausnitzii* (a butyrate producer with additional anti-inflammatory properties [Lopez-Siles et al., 2017]),

Bifidobacterium longum, and *Bifidobacterium breve* (two prevalent human bifidobacteria with immune-modulatory functions) (Henrick et al., 2021; Wong et al., 2020), and non-significantly with decreased gene richness ($p = 0.06$) (Table S1F). Thus, *F. prausnitzii*, *B. longum*, and *B. breve* might be markers of stable communities and/or possible stabilizing factors of the gut microbiota in our population.

To characterize the interactions between these bacteria and other community members, we performed co-occurrence network analysis on the longitudinal data using averages of repeated measurements, as described previously (Poyet et al., 2019). We observed that *B. longum* and *B. breve* interacted positively with each other, whereas *F. prausnitzii* was part of a more complex network and interacted positively with species from Ruminococcaceae, Lachnospiraceae, and Eubacteriaceae, such as the fiber degraders *Clostridium phytofermentans* and *Eubacterium cellulosolvens*, and the butyrate producer *Roseburia intestinalis* (Figure 2A; Table S2). These results indicate that *B. longum* and *B. breve* share a niche within the intestinal environment and that there might be different factors influencing the abundance of *B. longum*/*B. breve* and *F. prausnitzii* in the human gut. Human genetics (Kurilshikov et al., 2021; Schmidt et al., 2020) and the potential to degrade dietary and host-derived carbohydrates (Wong et al., 2020) influence the fecal abundance of *Bifidobacterium* species. Therefore, considering the lack of connections in our co-occurrence analysis, we hypothesize that the fecal abundance of *B. longum* and *B. breve* might be directly modulated by human genetics and diet, and possibly these bacteria might influence community stability indirectly through their anti-inflammatory properties and interaction with the immune system.

We noted that individuals with a high compositional variation often had large ranges of *F. prausnitzii*, *B. longum*, and *B. breve* abundance and/or a large variation of gene richness, with relatively low values at some of the sampling time points (Figure 2B). We observed a number of species whose change in abundance between visits was positively correlated with changes in *F. prausnitzii*, *B. longum*, or *B. breve* abundance and with changes in composition between visits (Figure 2C). The changes in *B. longum* and *B. breve* abundance correlated positively with each other and with abundance changes of other *Bifidobacterium* and Actinobacteria species, the lactate utilizer *Megasphaera* genomosp. type_1 and the butyrate producer *Coprococcus comes*, but not with abundance changes of *F. prausnitzii* (Figure 2C). The change in *F. prausnitzii* abundance correlated positively with the abundance change of several fiber degraders and butyrate producers (e.g., *Ruminococcus flavefaciens*, *C. phytofermentans*, *E. cellulosolvens*, *R. intestinalis*, and *Roseburia inulinivorans* [Louis and Flint, 2017]); the change in abundance of these species also correlated positively with changes in gene richness (Figure 2C).

Butyrate production is important for both microbial and host physiology (Koh et al., 2016; Louis and Flint, 2009): butyrate is not only the primary energy substrate for colonocytes but can also act as a histone deacetylase inhibitor, a ligand for G-protein-coupled receptors, and an immune-modulator to decrease inflammation and improve gut barrier function. Butyrate is produced as an end-product of colonic fermentation of fibers and carbohydrates that are not metabolized by the host, as well as from amino acids (Bui and de Vos, 2021). Two terminal enzymes are known for butyrate production from carbohydrates by the human gut microbiota: the butyryl CoA:acetate CoA transferase (*but*, the most abundant enzyme) and the butyrate kinase (*buk*, present in few strains). The terminal enzymes for butyrate production from amino acids are the butyryl CoA:4-hydroxybutyrate CoA transferase (*4-hbt*) and butyryl CoA:acetoacetate CoA transferase (*ato*), but their abundance in the human gut is lower than that of *but* and *buk* (Vital et al., 2014). We quantified the terminal genes for butyrate production and tested whether the microbial potential for butyrate production was linked to the intra-individual compositional variability (supplemental information; Figure S1; Table S1G). The abundance of *but* (the dominant terminal gene and exclusive gene for butyrate production in both *F. prausnitzii* and *Roseburia* species [Louis et al., 2010]) increased with increasing gene richness but was not linked to compositional variability. The abundance of *buk* increased with decreasing abundance of *F. prausnitzii* and with increasing abundance of *B. longum* and *B. breve*, consistent with the covariation of these bifidobacteria with *C. comes* (one of the few gut butyrate producers harboring *buk*; Louis and Flint, 2009) (Figure 2C). The abundance of *ato* increased with decreasing abundance of *F. prausnitzii* but also with increasing intra-individual compositional variability (Table S1G). These results suggest that compositional variability might be higher in conditions that favor butyrate production by *ato* (Figure S1), which might be determined not only by the substrate (carbohydrates versus proteins) but also by intestinal conditions, such as pH and redox potential, and the levels of acetate and lactate (Louis and Flint, 2017).

Finally, we observed a number of species (e.g., *Citrobacter*, *Escherichia*, and *Salmonella*) whose change in abundance between visits correlated negatively with changes in *F. prausnitzii* abundance (Figure 2C), and that showed negative correlations with the network of Firmicutes in the co-occurrence analysis (Figure 2A). The change in the abundance of *Escherichia* and *Salmonella* species did not correlate with intra-individual compositional variability (Figure 2C), and for *Escherichia coli*, we observed large abundance ranges for individuals with both large and small compositional variability (Figure 2B). These results suggest that the large variation of *E. coli* is a feature of the normal gut microbiota independent of community stability.

Patterns of total, inter-, and intra-individual variations for microbial species abundance

To determine the extent of total, intra-, and inter-individual variance for the abundance of the 384 microbial species identified in our cohort, we used mixed-effects models on the repeated fecal microbiota measurements (Table S3A). Given that knowledge of abundance ranges for key microbial taxa is considered a main step for the definition of a normal human gut microbiota (McBurney et al., 2019), we also calculated the central 95% reference interval for the abundance of the 384 species. To determine how much of the observed variance was due to variations between or within individuals, we calculated the intraclass correlation coefficient (ICC, defined as the ratio between the inter-individual and the total variation of the relative abundance) for each microbial species (Table S3A).

CellPress
OPEN ACCESS

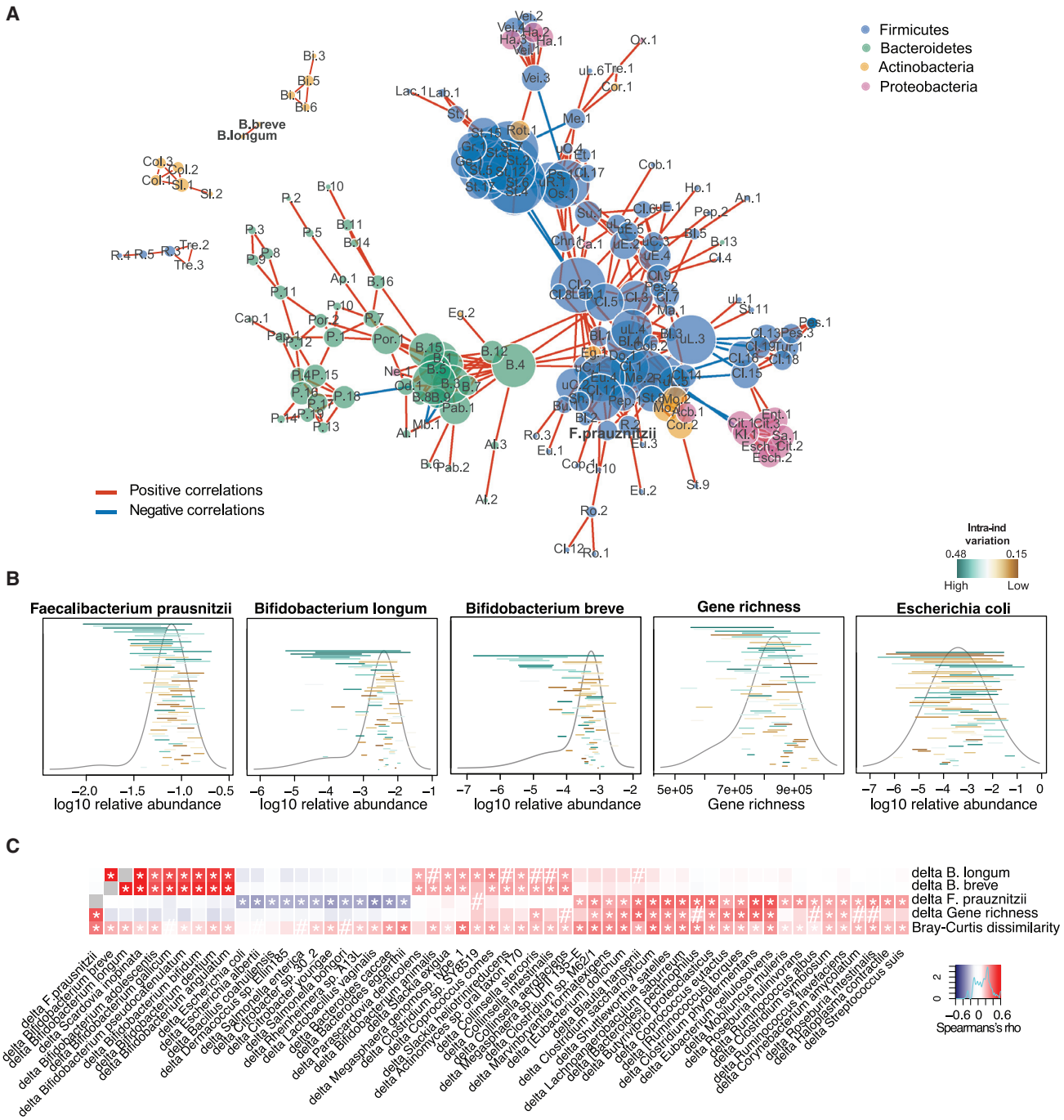


Figure 2. Species co-occurrence and covariation detected by longitudinal analysis

(A) Co-occurrence networks for microbial species in the S3WP cohort. Red and blue edges indicate significant positive and negative correlations, respectively. Nodes represent microbial species colored by phylum, with sizes proportional to the number of edges connected to the species (legend in Table S2).

(B) Histograms of the relative abundance of the indicated species and gene richness. Gray lines indicate sample frequency distributions. Horizontal bars indicate minimum-to-maximum ranges for the species' relative abundance in the four samples from each individual; bars are colored according to Bray-Curtis dissimilarity as shown in Figure 1D, with the brown to green scale indicating low to high gut microbiota compositional variability.

(C) Covariation of the abundance of *Faecalibacterium prausnitzii*, *Bifidobacterium longum*, and *Bifidobacterium breve* with the abundance of other species. Shown are Spearman's correlations for changes between visits (delta) of species abundances, gene richness, and Bray-Curtis dissimilarity (absolute values, as only positive correlations are possible). The plot shows results that have rho > 0.25 or < −0.25 and adjusted p < 0.1 in at least one of the rows. #adjusted p < 0.1; *adjusted p < 0.05.

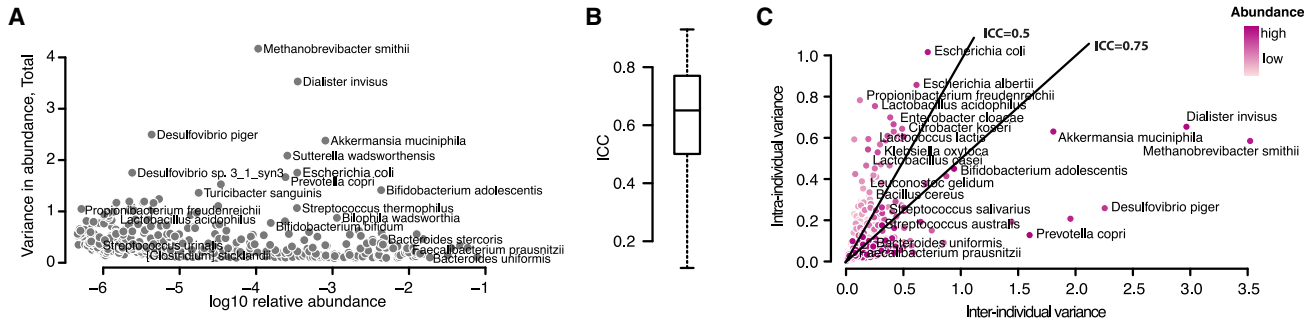


Figure 3. Patterns of inter- and intra-individual variation for microbial species

(A) Total variance of individual microbial species in the S3WP cohort, plotted against the relative abundance.

(B) Boxplot of the intraclass correlation coefficients (ICC) for the 384 prevalent gut microbiota species. Shown is the median and interquartile range (IQR); whiskers represent ±1.5 * IQR from the box's quartiles.

(C) Relationship between the inter-individual and intra-individual components of total species variance. Lines at ICC 0.5 and 0.75 separate species with high and low intra-individual variance, respectively. Dots indicate microbial species and are colored according to species abundance.

We observed that the total variance of abundance was negatively correlated with the relative abundance of microbial species (Spearman's correlation, rho = −0.30, p = 1.27e−09; Figure 3A). This finding is consistent with an earlier study (Mehta et al., 2018) and indicates that the abundance of dominant species is generally more stable. However, we observed a number of low-abundance species with low total variance (e.g., *C. phytofermentans*, *Megasphaera* genomosp. type_1, and the bile acid transforming *Clostridium scindens*) (Figure S2A; Table S3A), which are specialized taxa in gut communities. Ranges of abundance were one to two orders of magnitude for species with low total variance of abundance and up to five orders of magnitude for species with high total variance (Table S3A). The median ICC was 0.65 (Figure 3B), indicating that the inter-individual variation of species abundance was larger than the intra-individual variation for the majority of the 384 prevalent species.

We observed the lowest total variance for several abundant Firmicutes (e.g., *Clostridium* sp. *M62/1*, *Clostridium symbiosum*, *butyrate-producing bacterium SS3/4*, *Holdemania filiformis*, and *Subdoligranulum variabile*), which were generally more variable between individuals than in a person over time (ICC > 0.5; Table S3A). Dominant species such as *Ruminococcus obeum*, *C. comes*, *F. prausnitzii*, and *R. inulinivorans* (Figure S2A; Table S3A) also showed low total variance but had ICCs of around 0.5, indicating similar intra- and inter-individual variation and potential intra-individual bias.

A number of dominant Bacteroidetes species, such as *Bacteroides* sp. *9_1_42FAA*, *Bacteroides uniformis*, and *Alistipes putredinis*, had intermediate total variance and ICCs of around 0.75, indicating low intra-individual bias (Table S3A). A similar pattern was observed for *B. longum* and *B. breve* (Figure 3C; Table S3A).

We observed the largest total variance for *Methanobrevibacter smithii*, *Dialister invisus*, *Desulfovibrio piger*, *Akkermansia muciniphila*, *Sutterella wadsworthensis*, *Prevotella copri*, *Desulfovibrio* sp. *3_1_syn3*, and *E. coli* (Figures 3A and S2A; Table S3A). Among these species, *M. smithii*, *D. invisus*, *D. piger*, *A. muciniphila*, *S. wadsworthensis*, *P. copri*, and *D.* sp. *3_1_syn3* had ICCs ≈ 0.75 (Figure 3C; Table S3A), indicating high variability between individuals but low intra-individual bias. However, *E. coli* and *Escherichia albertii* had ICCs <0.5 (Figure 3C; Table S3A), indicating high intra-individual variation. Similarly, several species in Enterobacteriaceae (e.g., species in *Escherichia*, *Enterobacter*, and *Citrobacter*) and Lactobacillales (e.g., species in *Lactobacillus*, *Streptococcus*, *Lactococcus*, and *Leuconostoc*) showed large total variance due to intra-individual variation (Figure 3C; Table S3A); this pattern was also observed at the family level (Table S3B). Similar to our observations for *E. coli* (Figure 2B), the abundance of *Lactobacillus acidophilus* varied extensively between visits in the microbiota of individuals with both high and low compositional variability (Figure S2B; Tables S3A and S3B), suggesting that occasional blooming of these species is a normal feature of the human gut microbiota. Thus, reliable determination of abnormal abundances might be challenging for species with large intra-individual variation, and study designs with repeated sampling or observation of large effect sizes in large samples might be required to draw reliable conclusions for associations with health or disease states.

Similar results obtained using 16S rRNA gene profiling

As 16S rRNA gene profiling is a widely used method for gut microbiota studies, we repeated our analyses on 16S rRNA gene profiles in the V4 region using taxonomic information at genus level to test the robustness of our findings obtained using metagenomics. The results obtained from 16S rRNA gene analyses were consistent with those obtained from whole-genome metagenomics, and we observed similar patterns for the total intra- and inter-individual variability of microbial taxa (Figure S3; Table S3C). Our results are also consistent with a study using a phylogenetic microarray (i.e., the human intestinal tract chip or HITChip) showing stability over 8–12 years for the abundance of dominant Firmicutes (Rajilić-Stojanović et al., 2013) and with a 16S rRNA gene survey indicating *Akkermansia* and *Methanobrevibacter* among the main genera contributing to inter-individual variation in a cross-sectional study with 1,106 individuals (Falony et al., 2016).

Validation of results in an independent cohort

To validate our results in an independent cohort, we analyzed sequence data of three fecal samples from each of the 62 individuals from the HMP1-II cohort, a population cohort without

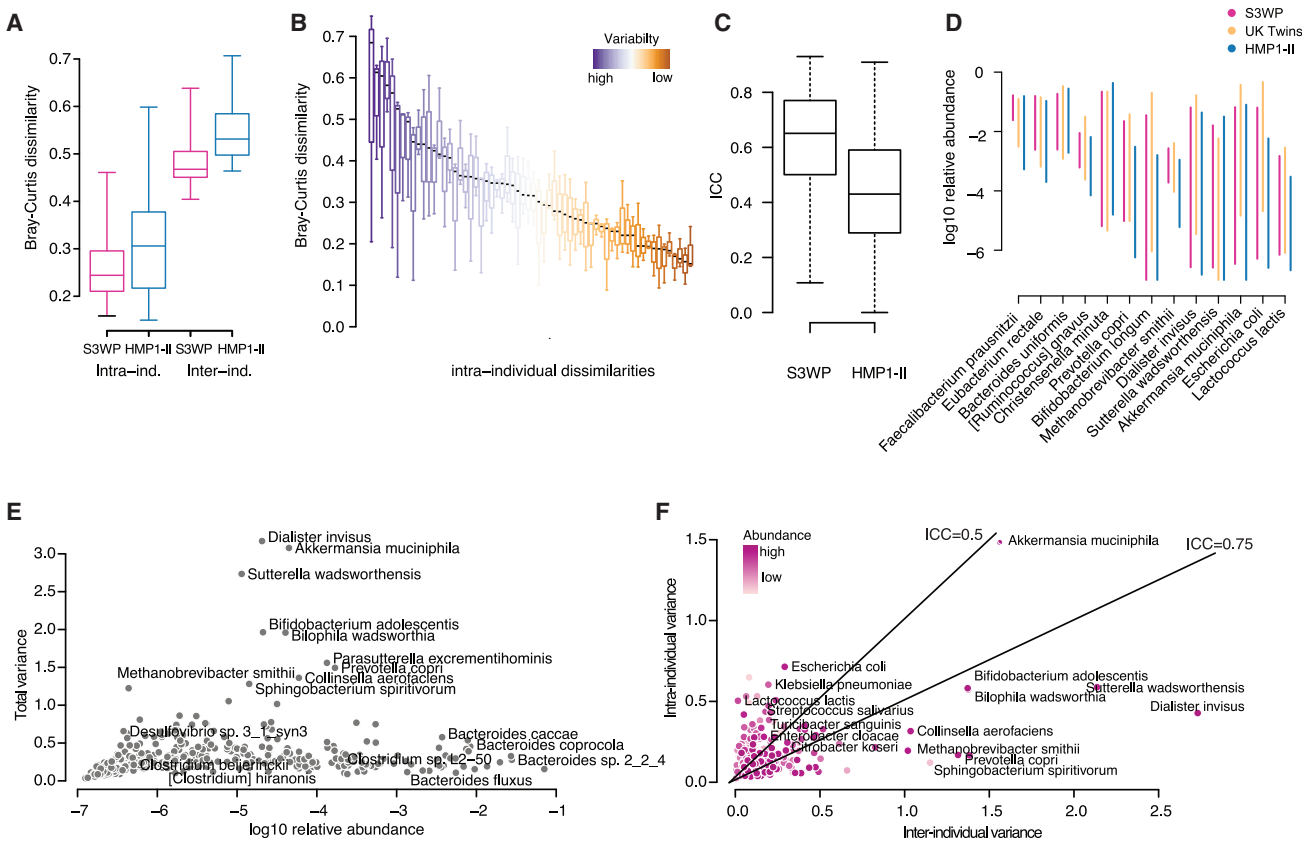


Figure 4. Validation of intra- and inter-individual variation patterns in the extended Human Microbiome Project

(A) Boxplot of average intra- and inter-individual compositional variability using Bray-Curtis dissimilarity for the S3WP and HMP1-II cohorts.

(B) Compositional variability of the gut microbiota for each of the 62 individuals in the HMP1-II cohort, estimated by the Bray-Curtis dissimilarity. Boxplots are ordered by the median intra-individual Bray-Curtis dissimilarity calculated for pairwise comparisons of the three samples.

(C) Boxplots of intraclass correlation coefficients (ICCs) for the S3WP and HMP1-II cohorts.

(D) Reference intervals for the abundance of selected microbial species, determined as the central 95% of the population, in the S3WP, UK Twins, and HMP1-II cohorts.

(E) Total variance of individual microbial species plotted against the relative abundance.

(F) Relationship between the inter-individual and intra-individual components of total species variance in the HMP1-II cohort. Lines at ICC 0.5 and 0.75 separate species with high and low intra-individual variance, respectively. Dots indicate microbial species and are colored according to species abundance. Boxes in (A)–(C) show median and interquartile ranges (IQRs); whiskers represent ±1.5 * IQR from each box's quartiles.

disease, aged 26.4 ± 5.1 years, and sampled 1–15 months apart (Lloyd-Price et al., 2017). The median compositional variation was larger between individuals than within individuals of the HMP1-II cohort (Figure 4A), and the degree of compositional variability varied between the individuals (Figure 4B), in line with results from the S3WP population (Figures 1B and 1D). We used a linear mixed-effect model on the 384 species identified in the S3WP cohort and did not observe associations between the abundance of any of these individual microbial species and gut microbiota stability in the HMP1-II cohort (Table S4). However, we confirmed that the inter- and intra-individual contributions to the variability of species abundances differed for different bacteria (Figure 4C).

The 20 most dominant species in the HMP1-II population included Bacteroidetes from the genera *Bacteroides*, *Parabacteroides*, and *Alistipes*, as well as *F. prausnitzii* (Table S4). Overall, the Bacteroidetes species had similar relative abundances, ranges of abundance, and ICC values in the HMP1-II and S3WP populations (Tables S3A and S4), but the ranges in the HMP1-II population were larger for *F. prausnitzii*, *E. rectale* and other Clostridiales (e.g., *R. gnavus* and *Christensenella minuta*), and *B. longum* (Figure 4D; Tables S3A and S4). These differences could at least partly be explained by the use of different DNA extraction protocols for analysis of the two populations. Repeated bead-beating, which was used to analyze the S3WP but not the HMP1-II population (Aagaard et al., 2013), might be required for effective extraction of microbial DNA from stools, particularly from Clostridia, whereas Bacteroidetes are less sensitive (Costea et al., 2017; Salonen et al., 2010).

D. invisus, *A. muciniphila*, *S. wadsworthensis*, and *P. copri* were among the top 10 species with the highest total variance in the HMP1-II cohort (Figure 4E; Table S4). These species were also highly variable in the S3WP cohort, with similar ranges of abundance (Figure 4D) and similar ICC values (ICC > 0.5) (Tables S3A and S4), indicating high inter-individual variation in both cohorts. The total variance of *M. smithii* was also high in

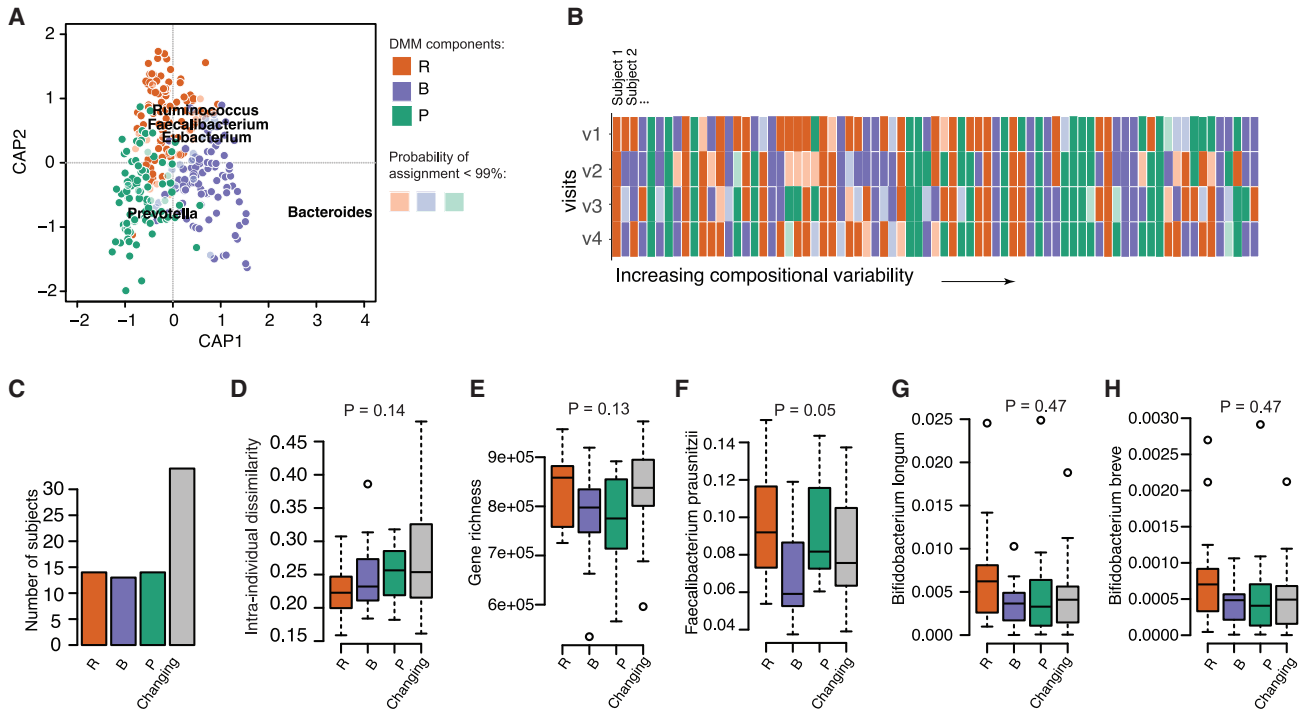


Figure 5. Stability of the enterotypes in the normal gut microbiota

(A) Genus-level microbiota community variation (two first components of constrained analysis of principal coordinate analysis [CAP], Bray-Curtis dissimilarity), constrained by the community types (enterotypes identified using Dirichlet multinomial mixtures [DMM]: R, *Ruminococcus*; B, *Bacteroides*; P, *Prevotella*). The 300 samples from the 75 individuals are colored according to the assigned community type. Shaded colors indicate samples with low certainty of the community type assignment (<99%).

(B) Enterotypes for each individual at the four visit times. Samples are colored according to the assigned community type; shaded colors indicate samples with low certainty of community type assignment (<99%).

(C) Histograms of the number of individuals with a stable enterotype at the four visits or a variable (changing) enterotype (gray bar). An individual is classified as changing enterotype based on assigned community types with a certainty higher than 99%.

(D–H) Compositional variability, gene richness, and abundance of the indicated species for individuals with stable or changing enterotype (p values from Kruskall-Wallis test). Boxes in (D)–(H) show median and interquartile ranges (IQRs); whiskers represent ±1.5 * IQR from each box's quartiles.

the HMP1-II cohort but markedly lower than in the S3WP cohort (as reflected in the ranges of abundance; Figure 4D). This difference could be explained by the lower abundance and prevalence of *M. smithii* in the HMP1-II samples (Table S4) or by differences in extraction protocols (Salonen et al., 2010), as indicated above.

Enterobacteriaceae, such as *E. coli*, *Enterobacter cloacae* and *Klebsiella oxytoca*, and Lactobacillales in the genera *Lactobacillus*, *Lactococcus*, and *Streptococcus* had high total and intra-individual variance in both cohorts (Figures 3C, 4E, and 4F; Tables S3A and S4). We also found consistent ranges of abundance in the UK Twins cohort (Figure 4D; Table S4), including samples from 1,004 individuals of a similar age to S3WP (65.0 ± 7.8 and 57.1 ± 4.1 years for UK Twins and S3WP, respectively) and DNA extracted with a bead-beating method (Visconti et al., 2019). These results validate the observations made for the S3WP cohort and show that a large intra-individual variation of Enterobacteriaceae is a feature of the normal gut microbiota, independent of regional and other potential differences, such as age.

Stability of enterotypes of the human gut microbiota

The human gut microbiota has been stratified into community types, or enterotypes, based on profiles of genera abundances (Arumugam et al., 2011; Ding and Schloss, 2014; Vieira-Silva et al., 2019). Although the enterotype classification has the potential for clinical applications, its value has been challenged (Costea et al., 2018; Gorvitovskaia et al., 2016; Jeffery et al., 2012; Knights et al., 2014). The stability of an individual's enterotype is a necessary condition for the use of enterotypes as predictive markers of health or disease. Therefore, we investigated the stability of enterotypes in the S3WP population at the four visits over the course of one year.

For the 300 samples of the 75 individuals, using Dirichlet multinomial mixtures (Holmes et al., 2012) we identified three community types, which were dominated by the genera *Ruminococcus*, *Bacteroides*, and *Prevotella*, consistent with the original definitions of enterotypes (Arumugam et al., 2011). Stratification by enterotypes explained 21% of the compositional variation in our dataset (distance-based redundancy analysis, dbRDA; Figure 5A).

We assigned enterotypes to each of the four samples from each individual (Figure 5B) and, as the discrete nature of the enterotypes has been challenged in favor of gradients (Koren et al., 2013), we sought to reduce uncertainties by disregarding samples that showed a low probability of assignment (certainty of assignment < 99%). Despite this permissive condition, 45%

CellPress
OPEN ACCESS

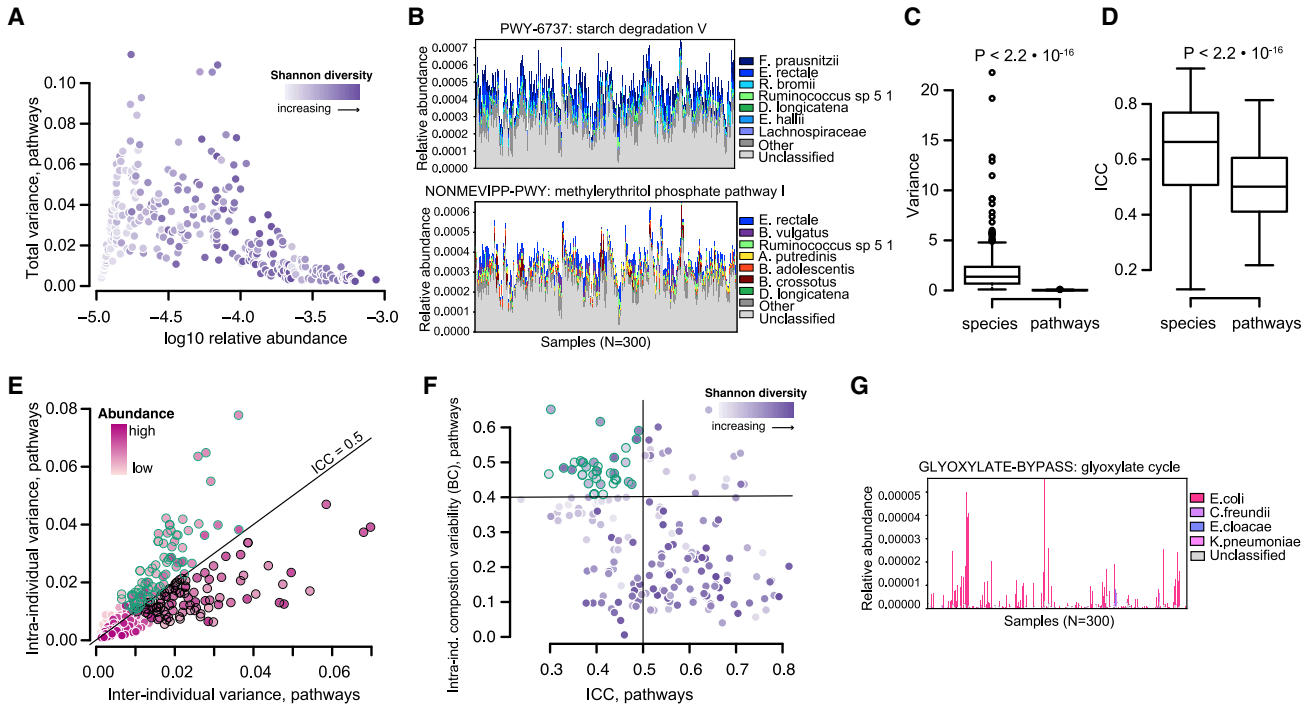


Figure 6. Intra-individual and inter-individual variation of the gut microbiota functional potential

(A) Total variance of pathways plotted against the relative abundance. Dots indicate pathways colored by the Shannon diversity index of their species assignments.

(B) Relative abundance of low variable gene families (starch degradation V [PWY-6737] [top] and methylerythritol phosphate pathway I [NONMEVIPP-PWY] [bottom]) in the 300 samples ordered by individual and visit.

(C) Boxplot of the total variance of microbial species and functional pathways.

(D) Boxplot of the intraclass correlation coefficients (ICC) for species and pathways.

(E) Relationship between inter- and intra-individual components of total variance in gene families. The line at ICC 0.5 separates gene families with high and low intra-individual variance. Dots circled in black indicate gene families with low variability (total variance < 0.01); dots circled in green indicate gene families with high intra-individual variability (total variance > 0.02 and ICC > 0.5).

(F) Pathway ICCs plotted against the intra-individual variability of the pathways' species assignments, calculated as the median Bray-Curtis (BC) dissimilarity. Dots indicate pathways colored by the Shannon diversity index of their species assignments.

(G) Relative abundance of the highly variable pathway "GLYOXYLATE-BYPASS: glyoxylate cycle" in the 300 samples ordered by individual and visit. Boxes in (C) and (D) show median and interquartile ranges (IQR); whiskers represent ±1.5 * IQR from each box's quartiles.

(34) of the 75 individuals changed their enterotype over the course of the study (Figure 5C [gray bar]). The remaining individuals kept the same enterotype over the one-year study period and were evenly divided between the *Ruminococcus*, *Bacteroides*, and *Prevotella* enterotypes (Figure 5C). We found no significant difference for intra-individual compositional variability, gene richness, and abundance of *F. prausnitzii*, *B. longum*, and *B. breve* between people who changed their enterotype and those who had a stable enterotype (Figures 5D–5H). Our results indicate that change of enterotype often occurs in the normal gut microbiota, in both variable and stable communities, and are consistent with previous findings showing that enterotypes likely reflect the variation of diet, transit time, and stool consistency in normal populations (Jeffery et al., 2012; Roager et al., 2016; Vandeputte et al., 2016).

Disease states such as inflammatory bowel disease and obesity comorbidities have been shown to associate with enterotype configurations other than *Ruminococcus*, *Bacteroides*, and *Prevotella*, such as enterotypes with a high variance in community structure (Holmes et al., 2012) and the low-cell-count *Bacteroides* B2 enterotype (Vieira-Silva et al., 2020; Vieira-Silva et al., 2019). As we did not observe these alternative configurations in the S3WP population (Figure S4), we conclude that the *Ruminococcus*, *Bacteroides*, and *Prevotella* enterotypes are all likely associated with disease-free states.

Total variance for gut microbiota functional potential

Among individuals, the functional potential of the gut microbiota is more conserved than the taxonomic species profile (Turnbaugh et al., 2009). However, information on temporal dynamics of the functional potential is limited. To characterize the normal variation of the gut microbiota functional potential, we used HUMAnN2 (Franzosa et al., 2018) and the MetaCyc database (Caspi et al., 2018) to determine gene families summarized into functional pathways and the species potentially assigned to each pathway in the metagenomic data of the S3WP population.

Consistent with our observation for the gut microbiota species (Figure 3A), the total variance of pathways was negatively correlated with the abundance of the pathways (Spearman's correlation, rho = −0.42, p = 2.167e−15; Figure 6A; Table S5), indicating

that abundant pathways were stable. Then, we calculated the Shannon diversity index to estimate richness and evenness of the species assigned to each pathway and observed a negative correlation between Shannon diversity and total variance of pathway abundance (Spearman's correlation, rho = −0.48, p < 2.2e−16). Therefore, these results are consistent with the known functional redundancy of the human gut microbiota, as they indicate that abundant and stable pathways are present in several bacterial species belonging to different taxonomic groups.

We observed several essential microbial functions among the abundant pathways with low total variance (abundance > 0.0001% and total variance < 0.01), such as synthesis of structural components (peptidoglycan, phospholipids, and nucleotides); biosynthesis of nine essential amino acids (aromatic amino acids, branched-chain amino acids, threonine, lysine, and methionine), vitamins and cofactors (e.g., thiamine and coenzyme-A), and terpenoids; and functions for carbohydrate catabolism, such as glycolysis (both from glucose and as part of sucrose degradation), pentose phosphate pathways, and partial starch hydrolysis to maltodextrins (i.e., starch degradation V) (Table S5). As expected, these functions were represented in a large number of species with diverse taxonomy (Figure 6B; Table S5).

Inter- and intra-individual variations for gut microbiota functional potential

As expected, the total variance for the abundance of pathways was lower than that of microbial species (Figure 6C). However, the contribution of intra-individual variance to the total variance was greater for functional pathways than for microbial species (i.e., lower median ICC for pathways; Figure 6D).

Pathways with large inter-individual variation (total variance > 0.02 and ICC > 0.5; Figure 6E [dots circled with black]) were functions for lipid metabolism (including both biosynthesis and beta-oxidation), histidine degradation, lactose and galactose degradation, the mannitol cycle, sulfate assimilation, and methanogenesis from CO_2 and H_2 (Table S5). Several of these pathways probably followed the distribution of specific gut microbes with high inter-individual variation, such as *M. smithii* for methanogenesis and *A. muciniphila* for the pathway SO4ASSIM-PWY (sulfate reduction I, assimilatory) (Figure S5).

Among the pathways with large intra-individual variance (total variance > 0.02 and ICC < 0.5; Figure 6E [dots circled with green]), we observed functions for catabolism of sugars, fermentative processes (e.g., production of lactate, acetate, propionate, and butyrate), the glyoxylate cycle, tricarboxylic acid (TCA) cycle and modified TCA cycles, and biosynthesis of components for electron transfer chains (e.g., phylloquinol, menaquinones, and demethylmenaquinones). We also observed large intra-individual fluctuations in potential functions for the synthesis of amino acids (e.g., arginine, tyrosine, and tryptophan) and production of cofactors such as vitamin K, biotin, and folate (Table S5).

To determine whether the intra-individual fluctuations in the abundance of functions were related to changes in the composition of species assigned to the pathways, we estimated pathway-specific compositional variability between the four samples from each individual using the Bray-Curtis distance and analyzed the species compositional variability for each pathway in relation to the pathway's ICC. We observed negative correlations for the compositional variability of pathways' taxonomic assignments with both Shannon diversity (Pearson correlation, r = −0.55, p < 2.2e−16) and ICC values (Pearson correlation, r = −0.40, p = 3.8e−09; Figure 6F). In particular, we observed that the pathways characterized by large intra-individual variation (ICC < 0.5; Figure 6F [dots circled with green]) also had variable taxonomic composition (pathway-specific intra-individual Bray-Curtis distance > 0.4) and a lower-than-average Shannon diversity. These results indicate that microbial functions that vary to a great extent in an individual over time are likely represented by a low number of microbial species that occasionally bloom. Among these pathways, we observed functions involved in the biosynthesis of aromatic amino acids and cofactors, as well as the glyoxylate cycle (Figures 6G and S5; Table S5). The main contributors to these pathways were often Gammaproteobacteria, particularly Enterobacteriaceae such as *E. coli* and *Citrobacter freundii* (Figures 6G and S5).

Using linear mixed-effects models, we identified pathways linked to gut microbiota temporal stability (Table S5). Pathways with increasing abundance in stable communities had low as well as high total variance. Among pathways with low total variance, we found the biosynthesis of branched-chain and other essential amino acids (i.e., lysine, threonine, and methionine) and cofactor *S*-adenosyl methionine salvage. As expected from their low total variance, a large number of bacteria from different taxonomic groups contributed to these pathways, including species with low total variance (e.g., *F. prausnitzii*, *Ruminococcus obeum*, and *B. longum*) and high inter-individual variance (*A. muciniphila* and *M. smithii*) (Figure S6; Table S5).

Pathways with high total variance showing increasing abundance in stable communities included synthesis of vitamins B1 and B12 and catabolism of sucrose and sugar acids (e.g., glucuronate, galacturonate, and fructuronate; found in plant polysaccharides, gums, intestinal mucus, and as detoxification products in the bile). A small number of bacteria with low total variance (e.g., *F. prausnitzii*) as well as high intra-individual variance (e.g., *E. coli*) contributed to these pathways (Figure S6; Table S5), in line with previous studies showing contribution of both *F. prausnitzii* and *E. coli* to the transcription of genes for galacturonate and glucuronate degradation (Schirmer et al., 2018). These results indicate that stable microbial communities have a high potential for biosynthesis of essential nutrients and utilization of carbon sources derived from both diet and the host. These results are potentially relevant for human health, as several studies have found decreased abundance of microbial genetic potential for metabolism of vitamins and cofactors in obesity, metabolic syndrome, and type 2 diabetes (Belda et al., 2022; Karlsson et al., 2013; Le Chatelier et al., 2013; Qin et al., 2012; Wu et al., 2020).

We also found several pathways with increasing abundance in less stable communities (Figure S6; Table S5). These pathways had low abundance and high intra-individual variation (represented among the pathways indicated by green circles in Figures 6F and 6G) and included catabolism of *N*-acetylneuraminic acid, glycerol, and sugar alcohols (e.g., mannitol, sorbitol, and galactitol) as well as the synthesis of menaquinones and redox components of electron transfer chains (Figure S6). The increase in these pathways in less stable communities possibly indicates a response to the availability of nutrients and electron

acceptors and altered redox conditions in the gut (Fischbach and Sonnenburg, 2011).

Our results overall indicate that high species redundancy in functional pathways reduces the variability and stabilizes the gut microbiota functional potential. However, our results also indicate that variable functions of the gut microbiota are represented by pathways with a large intra-individual component, which occur in a relatively small number of microbial species often belonging to Proteobacteria. We speculate that the temporal variation of some of these pathways might be important for gut microbiota homeostasis.

DISCUSSION

In this study, we investigated the temporal dynamics and the extent of intra- and inter-individual variation for the composition and functional potential of the normal gut microbiota. Our analyses showed that the intra-individual variation accounted for a large portion of the total variation in gut microbiota composition (23%) and that several functional pathways were highly dynamic (median ICC $\approx$ 0.5). However, we observed that the extent of intra-individual compositional variation was individual specific and lower in communities with high relative abundance of *F. prausnitzii*, *B. longum*, and *B. breve*, as well as with low temporal variation of *F. prausnitzii* and low abundance of the *ato* terminal gene for butyrate synthesis. We noted that all communities, both stable and variable, displayed occasional blooming of species and potential functions belonging to Enterobacteriaceae and other facultative anaerobic or aerotolerant bacteria. Therefore, our results suggest that these fluctuations are part of normal gut microbiota dynamics and homeostatic interaction with the host.

Cross-sectional studies have shown that known factors (e.g., host genetics, age, gut transit, diet, and medications) cumulatively explain about 20% of the overall gut microbiota compositional variation (Bonder et al., 2016; Falony et al., 2016; Goodrich et al., 2016; Rothschild et al., 2018; Wang et al., 2016). Our longitudinal analyses showed that the intra-individual variation accounted for a similar amount of variation, thus indicating that the intrinsic dynamics of gut microbiota should be considered in clinical studies as a source of bias when defining health and disease states. In addition, our observation that the gut microbiota composition of some individuals was highly variable suggests that community stability might not be sufficient to describe health in middle-aged populations such as the S3WP cohort.

Our finding of a link between the low abundance of *F. prausnitzii*, *B. longum*, and *B. breve* and high compositional variability offers insight into the potential mechanisms that influence gut microbiota community stability. *B. longum* and *B. breve* are closely related human-residential bifidobacteria that colonize the gut early in life but are still present in adults and have been described for their health promoting and anti-inflammatory functions (Arboleya et al., 2016; Wong et al., 2020). Similarly, the butyrate producer *F. prausnitzii* has anti-inflammatory properties and influences the gut's innate immune and barrier functions (Lopez-Siles et al., 2017) and might have additional roles for the recovery of gut communities after perturbation (Gibbons et al., 2017), such as response to changes in intestinal redox status using a flavin-thiol extracellular electron shuttle to scavenge oxygen (Khan et al., 2012). Interestingly, genetic variation of multiple riboflavin biosynthesis genes in *F. prausnitzii* has been linked to the number of leucocyte cells in blood (Chen et al., 2021), and oral supplementation of riboflavin in patients with Crohn's disease decreases systematic oxidative stress and the abundance of Enterobacteriaceae (von Martels et al., 2020). Therefore, our results might be clinically important, as low abundance and/or altered dynamics of *F. prausnitzii*, *B. longum*, and *B. breve* and high compositional variability might predict the risk of developing inflammatory intestinal and metabolic diseases, as indicated in recent studies (Clooney et al., 2021; Halfvarson et al., 2017; Reitmeier et al., 2020). However, we did not find a link between abundance of these species and gut microbiota compositional variability in the HMP1-II population. Given that the abundance of *F. prausnitzii*, *B. longum*, and *B. breve* declines with age (Arboleya et al., 2016; De Filippis et al., 2020) and that microbiota-based disease models are likely specific to geographic locations (He et al., 2018), our results highlight the importance of investigating localized and age-specific references for gut microbiota and their variability patterns.

We observed large fluctuations of *E. coli* and *L. acidophilus* in communities with both high and low compositional variability. Enterobacteriaceae have been identified in time-series analyses as conditionally rare taxa, which are able to occasionally bloom in response to disturbance (Gibbons et al., 2017). Temporal variability has also been reported for lactic acid bacteria, with *Lactobacillus* found to vary in the gut microbiota during circadian rhythm (Thaiss et al., 2014) and *Streptococcus* and *Lactococcus* described as conditionally rare taxa in different ecosystems (Lawson et al., 2015). Therefore, although the expansion of conditionally rare taxa might reflect the liberation of ecological niches in the gut (e.g., increased Enterobacteriaceae with intestinal inflammation), their fluctuation might reflect a dynamic response to perturbation or changed intestinal conditions, such as increased nutrient availability promoting fast growth (Song et al., 2017). Indeed, in line with our results in the S3WP population, the blooming of Proteobacteria and conditionally rare taxa has not been consistently linked to compositional shifts in the gut microbiota (Gibbons et al., 2017; Shade et al., 2014), and administration of *E. coli* and *Lactobacillus* spp. to an experimental model of neonatal dysbiosis has been shown to protect against late-onset sepsis and promote the establishment of normal gut homeostasis (Singer et al., 2019). In line with this observation and with a potential role in gut microbiota homeostasis, *E. coli* and *Lactobacillus* species are first colonizers of the human infant gut (Roswall et al., 2021) and bloom early during gut microbiota recovery after antibiotic treatment (Palleja et al., 2018). Therefore, current hypotheses proposing the expansion of Proteobacteria as the main diagnostic signature of gut microbiota dysbiosis and epithelial dysfunction (Litvak et al., 2017) should be carefully considered and validated by analyses of time-series in patient and control groups to control for possible different patterns of intra-individual dynamics. However, to assess the importance of variability patterns for community function, analyses of microbial activity, such as meta-transcriptomics and metabolomics, are also warranted, both in fecal and mucosal samples, as fecal sampling is a poor proxy for non-luminal microbes (Zoetendal et al., 2002).

Although our analyses confirm that the gut microbiota functional potential is more stable than species profiles (Turnbaugh et al., 2009), they also show that intra-individual variation might account for a larger portion of the total variance in functional pathways (about 50% based on the average ICC); thus, microbial functions also fluctuate in the normal gut microbiota over time. Cross-sectional studies show that functions with high inter-individual variability mostly belong to Proteobacteria as opposed to Firmicutes and Bacteroidetes (Bradley and Pollard, 2017). In the S3WP population, we observed that the functional pathways with large intra-individual variation belonged to a small number of Enterobacteriaceae, possibly suggesting that the functional variability observed cross-sectionally might be linked to the intra-individual dynamics of specific Enterobacteriaceae. The most variable functions were pathways for biosynthesis of aromatic amino acids and cofactors, which are essential nutrients for some symbionts, such as biotin and riboflavin for *F. prausnitzii* (Lopez-Siles et al., 2017). It is thus tempting to speculate that fluctuations in these pathways, possibly in response to altered dietary intake or intestinal bioavailability of nutrients, might reflect the metabolic status of both gut microbiota and the host. This hypothesis is consistent with our recent results showing that microbial biosynthesis of biotin is increased in individuals with prediabetes and diabetes (Wu et al., 2020).

In conclusion, our study underscores the importance of the temporal dynamics of gut microbiota for gut microbiome research and the need for localized longitudinal investigations in well-defined populations to assess the value of potential biomarkers for health or disease, particularly among features with high intra-individual bias.

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
- METHOD DETAILS
  - Study visits in the S3WP study
  - Extraction of total fecal genomic DNA
  - Metagenome sequencing of human fecal microbiota and analyses
  - 16S rRNA gene sequencing of human fecal microbiota
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.chom.2022.03.002.

ACKNOWLEDGMENTS

We thank Marie Louise Ekholm and Josefine Kjelldahl for study management and sample collection, Manuela Krämer and Robert Jakubowicz for metagenomics and 16S rRNA gene sequencing, and Eric Franzosa and Curtis Huttenhower for sharing data from the extended Human Microbiome Project. The computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX, partially funded by the Swedish Research Council through grant agreement no. 2016-07213. This study was supported by the Knut and Alice Wallenberg Foundation (2017.0026), JPI (A healthy diet for a healthy life; 2017-01996_3), Swedish Research Council (2019-01599), the Leducq Foundation (17CVD01), and grants from the Swedish state under the agreement between the Swedish government and the county councils, the ALF-agreement (ALFGBG-718101 and ALFGBG-718851). F. Bäckhed. is a recipient of a European Research Council (ERC) Consolidator Grant (615362; METABASE) and is the Torsten Söderberg Professor in Medicine and a Wallenberg Scholar.

AUTHOR CONTRIBUTIONS

M.U., G.B., and F. Bäckhed designed the study. G.B., A.G., and L.F. coordinated the study and collected the clinical data. L.O., V.T., F. Boulund, L.F., and L.E. coordinated the microbiome sample inventory and sequencing. L.O., F. Bäckhed, and V.T., conceptualized the analyses. L.O., F. Boulund, and S.N. performed data processing and analyses. F. Bäckhed, V.T., and S.N. supervised the project. M.U., G.B., L.E., and F. Bäckhed. secured the funding. L.O., V.T., R.P., and F. Bäckhed. interpreted the data and wrote the paper with input from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 19, 2021
Revised: January 17, 2022
Accepted: March 3, 2022
Published: March 28, 2022

REFERENCES

Aagaard, K., Petrosino, J., Keitel, W., Watson, M., Katancik, J., Garcia, N., Patel, S., Cutting, M., Madden, T., Hamilton, H., et al. (2013). The Human Microbiome Project strategy for comprehensive sampling of the human microbiome and why it matters. FASEB J. *27*, 1012–1022.

Arboleya, S., Watkins, C., Stanton, C., and Ross, R.P. (2016). Gut bifidobacteria populations in human health and aging. Front. Microbiol. *7*, 1204.

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., Fernandes, G.R., Tap, J., Bruls, T., Batto, J.M., et al. (2011). Enterotypes of the human gut microbiome. Nature *473*, 174–180.

Bäckhed, F., Fraser, C.M., Ringel, Y., Sanders, M.E., Sartor, R.B., Sherman, P.M., Versalovic, J., Young, V., and Finlay, B.B. (2012). Defining a healthy human gut microbiome: current concepts, future directions, and clinical applications. Cell Host Microbe *12*, 611–622.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. J. Stat. Soft. *67*, 1–48.

Belda, E., Voland, L., Tremaroli, V., Falony, G., Adriouch, S., Assmann, K.E., Prifiti, E., Aron-Wisnewsky, J., Debédat, J., Le Roy, T., et al. (2022). Impairment of gut microbial biotin metabolism and host biotin status in severe obesity: effect of biotin and prebiotic supplementation on improved metabolism. Gut. https://doi.org/10.1136/gutjnl-2021-325753.

Bergström, G., Berglund, G., Blomberg, A., Brandberg, J., Engström, G., Engvall, J., Eriksson, M., de Faire, U., Flinck, A., Hansson, M.G., et al. (2015). The Swedish CArdioPulmonary BioImage Study: objectives and design. J. Intern. Med. *278*, 645–659.

Bonder, M.J., Kurilshikov, A., Tigchelaar, E.F., Mujagic, Z., Imhann, F., Vila, A.V., Deelen, P., Vatanen, T., Schirmer, M., Smeekens, S.P., et al. (2016). The effect of host genetics on the gut microbiome. Nat. Genet. *48*, 1407–1412.

Bradley, P.H., and Pollard, K.S. (2017). Proteobacteria explain significant functional variability in the human gut microbiome. Microbiome *5*, 36.

Bui, T.P.N., and de Vos, W.M. (2021). Next-generation therapeutic bacteria for treatment of obesity, diabetes, and other endocrine diseases. Best Pract. Res. Clin. Endocrinol. Metab. *35*, 101504.

Caporaso, J.G., Bittinger, K., Bushman, F.D., DeSantis, T.Z., Andersen, G.L., and Knight, R. (2010a). PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics *26*, 266–267.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010b). QIIME allows analysis of high-throughput community sequencing data. Nat. Methods *7*, 335–336.

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*, D633–D639.

Chen, L., Wang, D., Garmaeva, S., Kurilshikov, A., Vich Vila, A., Gacesa, R., Sinha, T., Lifelines Cohort Study, Segal, E., Weersma, R.K., et al. (2021). The long-term genetic stability and individual specificity of the human gut microbiome. Cell *184*, 2302–2315.e2312.

Clooney, A.G., Eckenberger, J., Laserna-Mendieta, E., Sexton, K.A., Bernstein, M.T., Vagianos, K., Sargent, M., Ryan, F.J., Moran, C., Sheehan, D., et al. (2021). Ranking microbiome variance in inflammatory bowel disease: a large longitudinal intercontinental study. Gut *70*, 499–510.

Costea, P.I., Hildebrand, F., Arumugam, M., Bäckhed, F., Blaser, M.J., Bushman, F.D., de Vos, W.M., Ehrlich, S.D., Fraser, C.M., Hattori, M., et al. (2018). Enterotypes in the landscape of gut microbial community composition. Nat. Microbiol. *3*, 8–16.

Costea, P.I., Zeller, G., Sunagawa, S., Pelletier, E., Alberti, A., Levenez, F., Tramontano, M., Driessen, M., Hercog, R., Jung, F.E., et al. (2017). Towards standards for human fecal sample processing in metagenomic studies. Nat. Biotechnol. *35*, 1069–1076.

Costello, E.K., Lauber, C.L., Hamady, M., Fierer, N., Gordon, J.I., and Knight, R. (2009). Bacterial community variation in human body habitats across space and time. Science *326*, 1694–1697.

Csardi, G., and Nepusz, T. (2006). The igraph software package for complex network research. InterJournal Complex Syst. *1695*, 1–9.

David, L.A., Materna, A.C., Friedman, J., Campos-Baptista, M.I., Blackburn, M.C., Perrotta, A., Erdman, S.E., and Alm, E.J. (2014). Host lifestyle affects human microbiota on daily timescales. Genome Biol. *15*, R89.

De Filippis, F., Pasolli, E., and Ercolini, D. (2020). Newly explored Faecalibacterium diversity is connected to age, lifestyle, geography, and disease. Curr. Biol. *30*, 4932–4943.e4.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G.L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl. Environ. Microbiol. *72*, 5069–5072.

Ding, T., and Schloss, P.D. (2014). Dynamics and associations of microbial community types across the human body. Nature *509*, 357–360.

Duvallet, C., Gibbons, S.M., Gurry, T., Irizarry, R.A., and Alm, E.J. (2017). Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. Nat. Commun. *8*, 1784.

Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics *26*, 2460–2461.

Faith, D.P. (1994). Phylogenetic pattern and the quantification of organismal biodiversity. Philos. Trans. R. Soc. Lond. B Biol. Sci. *345*, 45–58.

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.

Fischbach, M.A., and Sonnenburg, J.L. (2011). Eating for two: how metabolism establishes interspecies interactions in the gut. Cell Host Microbe *10*, 336–347.

Flores, G.E., Caporaso, J.G., Henley, J.B., Rideout, J.R., Domogala, D., Chase, J., Leff, J.W., Vázquez-Baeza, Y., Gonzalez, A., Knight, R., et al. (2014). Temporal variability is a personalized feature of the human microbiome. Genome Biol. *15*, 531.

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.

Friedman, J., and Alm, E.J. (2012). Inferring correlation networks from genomic survey data. PLoS Comput. Biol. *8*, e1002687.

Gibbons, S.M., Kearney, S.M., Smillie, C.S., and Alm, E.J. (2017). Two dynamic regimes in the human gut microbiome. PLoS Comput. Biol. *13*, e1005364.

Goodrich, J.K., Davenport, E.R., Beaumont, M., Jackson, M.A., Knight, R., Ober, C., Spector, T.D., Bell, J.T., Clark, A.G., and Ley, R.E. (2016). Genetic determinants of the gut microbiome in UK Twins. Cell Host Microbe *19*, 731–743.

Gorvitovskaia, A., Holmes, S.P., and Huse, S.M. (2016). Interpreting Prevotella and Bacteroides as biomarkers of diet and lifestyle. Microbiome *4*, 15.

Haas, B.J., Gevers, D., Earl, A.M., Feldgarden, M., Ward, D.V., Giannoukos, G., Ciulla, D., Tabbaa, D., Highlander, S.K., Sodergren, E., et al. (2011). Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. Genome Res. *21*, 494–504.

Halfvarson, J., Brislawn, C.J., Lamendella, R., Vázquez-Baeza, Y., Walters, W.A., Bramer, L.M., D'Amato, M., Bonfiglio, F., McDonald, D., Gonzalez, A., et al. (2017). Dynamics of the human gut microbiome in inflammatory bowel disease. Nat. Microbiol. *2*, 17004.

He, Y., Wu, W., Zheng, H.M., Li, P., McDonald, D., Sheng, H.F., Chen, M.X., Chen, Z.H., Ji, G.Y., Zheng, Z.D., et al. (2018). Regional variation limits applications of healthy gut microbiome reference ranges and disease models. Nat. Med. *24*, 1532–1535.

Henrick, B.M., Rodriguez, L., Lakshmikanth, T., Pou, C., Henckel, E., Arzoomand, A., Olin, A., Wang, J., Mikes, J., Tan, Z., et al. (2021). Bifidobacteria-mediated immune system imprinting early in life. Cell *184*, 3884–3898.e11.

Holmes, I., Harris, K., and Quince, C. (2012). Dirichlet multinomial mixtures: generative models for microbial metagenomics. PLoS One *7*, e30126.

Human Microbiome Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207–214.

Jackson, M.A., Goodrich, J.K., Maxan, M.E., Freedberg, D.E., Abrams, J.A., Poole, A.C., Sutter, J.L., Welter, D., Ley, R.E., Bell, J.T., et al. (2016). Proton pump inhibitors alter the composition of the gut microbiota. Gut *65*, 749–756.

Jeffery, I.B., Claesson, M.J., O'Toole, P.W., and Shanahan, F. (2012). Categorization of the gut microbiota: enterotypes or gradients? Nat. Rev. Microbiol. *10*, 591–592.

Karlsson, F.H., Nookaew, I., and Nielsen, J. (2014). Metagenomic data utilization and analysis (MEDUSA) and construction of a global gut microbial gene catalogue. PLoS Comput. Biol. *10*, e1003706.

Karlsson, F.H., Tremaroli, V., Nookaew, I., Bergström, G., Behre, C.J., Fagerberg, B., Nielsen, J., and Bäckhed, F. (2013). Gut metagenome in European women with normal, impaired and diabetic glucose control. Nature *498*, 99–103.

Khan, M.T., Duncan, S.H., Stams, A.J., van Dijl, J.M., Flint, H.J., and Harmsen, H.J. (2012). The gut anaerobe Faecalibacterium prausnitzii uses an extracellular electron shuttle to grow at oxic-anoxic interphases. ISME J. *6*, 1578–1585.

Knights, D., Ward, T.L., McKinlay, C.E., Miller, H., Gonzalez, A., McDonald, D., and Knight, R. (2014). Rethinking "enterotypes.". Cell Host Microbe *16*, 433–437.

Koh, A., De Vadder, F., Kovatcheva-Datchary, P., and Bäckhed, F. (2016). From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. Cell *165*, 1332–1345.

Koren, O., Knights, D., Gonzalez, A., Waldron, L., Segata, N., Knight, R., Huttenhower, C., and Ley, R.E. (2013). A guide to enterotypes across the human body: meta-analysis of microbial community structures in human microbiome datasets. PLoS Comput. Biol. *9*, e1002863.

Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl. Environ. Microbiol. *79*, 5112–5120.

Kurilshikov, A., Medina-Gomez, C., Bacigalupe, R., Radjabzadeh, D., Wang, J., Demirkan, A., Le Roy, C.I., Raygoza Garay, J.A., Finnicum, C.T., Liu, X., et al. (2021). Large-scale association analyses identify host factors influencing human gut microbiome composition. Nat. Genet. *53*, 156–165.

Kurtz, Z.D., Müller, C.L., Miraldi, E.R., Littman, D.R., Blaser, M.J., and Bonneau, R.A. (2015). Sparse and compositionally robust inference of microbial ecological networks. PLoS Comput. Biol. *11*, e1004226.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lawson, C.E., Strachan, B.J., Hanson, N.W., Hahn, A.S., Hall, E.R., Rabinowitz, B., Mavinic, D.S., Ramey, W.D., and Hallam, S.J. (2015). Rare taxa have potential to make metabolic contributions in enhanced biological phosphorus removal ecosystems. Environ. Microbiol. *17*, 4979–4993.

Le Chatelier, E., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., Almeida, M., Arumugam, M., Batto, J.M., Kennedy, S., et al. (2013). Richness of human gut microbiome correlates with metabolic markers. Nature *500*, 541–546.

Li, J., Jia, H., Cai, X., Zhong, H., Feng, Q., Sunagawa, S., Arumugam, M., Kultima, J.R., Prifti, E., Nielsen, T., et al. (2014). An integrated catalog of reference genes in the human gut microbiome. Nat. Biotechnol. *32*, 834–841.

Litvak, Y., Byndloss, M.X., Tsolis, R.M., and Bäumler, A.J. (2017). Dysbiotic Proteobacteria expansion: a microbial signature of epithelial dysfunction. Curr. Opin. Microbiol. *39*, 1–6.

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.

Lopez-Siles, M., Duncan, S.H., Garcia-Gil, L.J., and Martinez-Medina, M. (2017). Faecalibacterium prausnitzii: from microbiology to diagnostics and prognostics. ISME J. *11*, 841–852.

Louis, P., and Flint, H.J. (2009). Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. FEMS Microbiol. Lett. *294*, 1–8.

Louis, P., and Flint, H.J. (2017). Formation of propionate and butyrate by the human colonic microbiota. Environ. Microbiol. *19*, 29–41.

Louis, P., Young, P., Holtrop, G., and Flint, H.J. (2010). Diversity of human colonic butyrate-producing bacteria revealed by analysis of the butyryl-CoA:acetate CoA-transferase gene. Environ. Microbiol. *12*, 304–314.

Lozupone, C., and Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. Appl. Environ. Microbiol. *71*, 8228–8235.

Lynch, S.V., and Pedersen, O. (2016). The human intestinal microbiome in health and disease. N. Engl. J. Med. *375*, 2369–2379.

Mardinoglu, A., Wu, H., Bjornson, E., Zhang, C., Hakkarainen, A., Räsänen, S.M., Lee, S., Mancina, R.M., Bergentall, M., Pietiläinen, K.H., et al. (2018). An integrated understanding of the rapid metabolic benefits of a carbohydrate-restricted diet on hepatic steatosis in humans. Cell Metab. *27*, 559–571.e5.

McBurney, M.I., Davis, C., Fraser, C.M., Schneeman, B.O., Huttenhower, C., Verbeke, K., Walter, J., and Latulippe, M.E. (2019). Establishing what constitutes a healthy human gut microbiome: state of the science, regulatory considerations, and future directions. J. Nutr. *149*, 1882–1895.

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.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, H., Wagner, S., et al. (2015). vegan: community ecology package. https://cran.r-project.org.

Palleja, A., Mikkelsen, K.H., Forslund, S.K., Kashani, A., Allin, K.H., Nielsen, T., Hansen, T.H., Liang, S., Feng, Q., Zhang, C., et al. (2018). Recovery of gut microbiota of healthy adults following antibiotic exposure. Nat. Microbiol. *3*, 1255–1265.

Poyet, M., Groussin, M., Gibbons, S.M., Avila-Pacheco, J., Jiang, X., Kearney, S.M., Perrotta, A.R., Berdy, B., Zhao, S., Lieberman, T.D., et al. (2019). A library of human gut bacterial isolates paired with longitudinal multiomics data enables mechanistic microbiome research. Nat. Med. *25*, 1442–1452.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS One *5*, e9490.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. Nature *464*, 59–65.

Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., Shen, D., et al. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature *490*, 55–60.

R Core Team (2018). R: a language and environment for statistical computing (R Foundation for Statistical Computing).

Rajilić-Stojanović, M., Heilig, H.G.H.J., Tims, S., Zoetendal, E.G., and de Vos, W.M. (2013). Long-term monitoring of the human intestinal microbiota composition. Environmental Microbiology *15*, 1146–1159.

Reitmeier, S., Kiessling, S., Clavel, T., List, M., Almeida, E.L., Ghosh, T.S., Neuhaus, K., Grallert, H., Linseisen, J., Skurk, T., et al. (2020). Arrhythmic gut microbiome signatures predict risk of type 2 diabetes. Cell Host Microbe *28*, 258–272.e6.

Roager, H.M., Hansen, L.B., Bahl, M.I., Frandsen, H.L., Carvalho, V., Gøbel, R.J., Dalgaard, M.D., Plichta, D.R., Sparholt, M.H., Vestergaard, H., et al. (2016). Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. Nat. Microbiol. *1*, 16093.

Roswall, J., Olsson, L.M., Kovatcheva-Datchary, P., Nilsson, S., Tremaroli, V., Simon, M.C., Kiilerich, P., Akrami, R., Krämer, M., Uhlén, M., et al. (2021). Developmental trajectory of the healthy human gut microbiota during the first 5 years of life. Cell Host Microbe *29*, 765–776.e3.

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.

Salonen, A., Lahti, L., Salojärvi, J., Holtrop, G., Korpela, K., Duncan, S.H., Date, P., Farquharson, F., Johnstone, A.M., Lobley, G.E., et al. (2014). Impact of diet and individual variation on intestinal microbiota composition and fermentation products in obese men. ISME J. *8*, 2218–2230.

Salonen, A., Nikkilä, J., Jalanka-Tuovinen, J., Immonen, O., Rajilić-Stojanović, M., Kekkonen, R.A., Palva, A., and de Vos, W.M. (2010). Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. J. Microbiol. Methods *81*, 127–134.

Schirmer, M., Franzosa, E.A., Lloyd-Price, J., McIver, L.J., Schwager, R., Poon, T.W., Ananthakrishnan, A.N., Andrews, E., Barron, G., Lake, K., et al. (2018). Dynamics of metatranscription in the inflammatory bowel disease gut microbiome. Nat. Microbiol. *3*, 337–346.

Schmidt, T.S.B., Raes, J., and Bork, P. (2018). The human gut microbiome: From association to modulation. Cell *172*, 1198–1215.

Schmidt, V., Enav, H., Spector, T.D., Youngblut, N.D., and Ley, R.E. (2020). Strain-level analysis of *Bifidobacterium* spp. from gut microbiomes of adults with differing lactase persistence genotypes. mSystems *5*, e00911–e00920.

Shade, A., Jones, S.E., Caporaso, J.G., Handelsman, J., Knight, R., Fierer, N., and Gilbert, J.A. (2014). Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. mBio *5*, e01314–e01371.

Singer, J.R., Blosser, E.G., Zindl, C.L., Silberger, D.J., Conlan, S., Laufer, V.A., DiToro, D., Deming, C., Kumar, R., Morrow, C.D., et al. (2019). Preventing dysbiosis of the neonatal mouse intestinal microbiome protects against late-onset sepsis. Nat. Med. *25*, 1772–1782.

Sinha, R., Abu-Ali, G., Vogtmann, E., Fodor, A.A., Ren, B., Amir, A., Schwager, E., Crabtree, J., Ma, S., et al.; Microbiome Quality Control Project Consortium (2017). Assessment of variation in microbial community amplicon sequencing

by the Microbiome Quality Control (MBQC) project consortium. Nat. Biotechnol. *35*, 1077–1086.

Song, H.K., Song, W., Kim, M., Tripathi, B.M., Kim, H., Jablonski, P., and Adams, J.M. (2017). Bacterial strategies along nutrient and time gradients, revealed by metagenomic analysis of laboratory microcosms. FEMS Microbiol. Ecol. *93*.

Storey, J.D., Bass, A.J., Dabney, A., and Robinson, D. (2015). qvalue: Q-value estimation for false discovery rate control. R package version 2.16.0. http://github.com/jdstorey/qvalue.

Sze, M.A., and Schloss, P.D. (2016). Looking for a signal in the noise: revisiting obesity and the microbiome. mBio *7*, e01018–16.

Tebani, A., Gummesson, A., Zhong, W., Koistinen, I.S., Lakshmikanth, T., Olsson, L.M., Boulund, F., Neiman, M., Stenlund, H., Hellström, C., et al. (2020). Integration of molecular profiles in a longitudinal wellness profiling cohort. Nat. Commun. *11*, 4487.

Thaiss, C.A., Zeevi, D., Levy, M., Zilberman-Schapira, G., Suez, J., Tengeler, A.C., Abramson, L., Katz, M.N., Korem, T., Zmora, N., et al. (2014). Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. Cell *159*, 514–529.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P., et al. (2009). A core gut microbiome in obese and lean twins. Nature *457*, 480–484.

Valdes, A.M., Walter, J., Segal, E., and Spector, T.D. (2018). Role of the gut microbiota in nutrition and health. BMJ *361*, k2179.

Vandeputte, D., Falony, G., Vieira-Silva, S., Tito, R.Y., Joossens, M., and Raes, J. (2016). Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. Gut *65*, 57–62.

Vieira-Silva, S., Falony, G., Belda, E., Nielsen, T., Aron-Wisnewsky, J., Chakaroun, R., Forslund, S.K., Assmann, K., Valles-Colomer, M., Nguyen, T.T.D., et al. (2020). Statin therapy is associated with lower prevalence of gut microbiota dysbiosis. Nature *581*, 310–315.

Vieira-Silva, S., Sabino, J., Valles-Colomer, M., Falony, G., Kathagen, G., Caenepeel, C., Cleynen, I., van der Merwe, S., Vermeire, S., and Raes, J. (2019). Quantitative microbiome profiling disentangles inflammation- and bile duct obstruction-associated microbiota alterations across PSC/IBD diagnoses. Nat. Microbiol. *4*, 1826–1831.

Visconti, A., Le Roy, C.I., Rosa, F., Rossi, N., Martin, T.C., Mohney, R.P., Li, W., de Rinaldis, E., Bell, J.T., Venter, J.C., et al. (2019). Interplay between the human gut microbiome and host metabolism. Nat. Commun. *10*, 4505.

Vital, M., Howe, A.C., and Tiedje, J.M. (2014). Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. mBio *5*, e00889.

von Martels, J.Z.H., Bourgonje, A.R., Klaassen, M.A.Y., Alkhalifah, H.A.A., Sadaghian Sadabad, M., Vich Vila, A., Gacesa, R., Gabriëls, R.Y., Steinert, R.E., Jansen, B.H., et al. (2020). Riboflavin supplementation in patients with Crohn's disease [the RISE-UP study]. J. Crohns Colitis *14*, 595–607.

Wang, J., Thingholm, L.B., Skiecevičienė, J., Rausch, P., Kummen, M., Hov, J.R., Degenhardt, F., Heinsen, F.A., Rühlemann, M.C., Szymczak, S., et al. (2016). Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. Nat. Genet. *48*, 1396–1406.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microbiol. *73*, 5261–5267.

Wong, C.B., Odamaki, T., and Xiao, J.Z. (2020). Insights into the reason of Human-Residential Bifidobacteria (HRB) being the natural inhabitants of the human gut and their potential health-promoting benefits. FEMS Microbiol. Rev. *44*, 369–385.

Wu, H., Esteve, E., Tremaroli, V., Khan, M.T., Caesar, R., Mannerås-Holm, L., Ståhlman, M., Olsson, L.M., Serino, M., Planas-Fèlix, M., et al. (2017). Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug. Nat. Med. *23*, 850–858.

Wu, H., Tremaroli, V., Schmidt, C., Lundqvist, A., Olsson, L.M., Krämer, M., Gummesson, A., Perkins, R., Bergström, G., and Bäckhed, F. (2020). The gut microbiota in prediabetes and diabetes: A population-based cross-sectional study. Cell Metab. *32*, 379–390.e3.

Xavier, R.J. (2016). Microbiota as therapeutic targets. Dig. Dis. *34*, 558–565.

Zaneveld, J.R., McMinds, R., and Vega Thurber, R. (2017). Stress and stability: applying the Anna Karenina principle to animal microbiomes. Nat. Microbiol. *2*, 17121.

Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics *30*, 614–620.

Zoetendal, E.G., von Wright, A., Vilpponen-Salmela, T., Ben-Amor, K., Akkermans, A.D., and de Vos, W.M. (2002). Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. Appl. Environ. Microbiol. *68*, 3401–3407.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Human feces	This paper	N/A
Critical commercial assays		
TruSeq DNA PCRfree library preparation kit	Illumina	Cat# 20015963
5PRIME HotMasterMix	Quanta bio	Cat# 2200410
NucleoSpin Gel and PCR Clean-up kit	MACHEREY-NAGEL	Cat# 740609
Quant-iT PicoGreen dsDNA kit	Thermo Fischer	Cat# P11496
MiSeq Reagent Kit v2 (500-cycles)	Illumina	Cat# MS-102-2003
Deposited data		
Whole genome metagenomic sequencing data	This paper	ENA: PRJEB38984
16S rRNA sequencing data	This paper	ENA: PRJEB38984
Oligonucleotides		
16S rRNA sequencing primers	(Kozich et al., 2013)	See Table S6
Software and algorithms		
R Software	(R Core Team, 2018)	https://www.r-project.org
MEDUSA	(Karlsson et al., 2014)	N/A
HUMANN2	(Franzosa et al., 2018)	https://huttenhower.sph.harvard.edu/humann2/
DMM	(Holmes et al., 2012)	https://bioconductor.org/packages/3.14/bioc/html/DirichletMultinomial.html
Qiime (v1.9)	(Caporaso et al., 2010b)	http://qiime.org
Other		
FastPrep-24 Instrument	MP Biomedicals	Cat# 116004500

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Fredrik Bäckhed (fredrik@wlab.gu.se).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Whole-genome metagenomic and 16S rRNA gene sequencing data analyzed in this study have been deposited to the European Nucleotide Archive (ENA) under the study accession PRJEB38984 and is publicly available.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

101 individuals were recruited to the Swedish SciLifeLab SCAPIS Wellness Profiling (S3WP) study from SCAPIS, a prospective observational study of men and women aged 50-65 years randomly selected from the general Swedish population (Bergström et al., 2015). Examinations in SCAPIS include imaging to assess coronary and carotid atherosclerosis, clinical chemistry, anthropometry, and extensive questionnaires (Bergström et al., 2015). No exclusion criteria were applied in SCAPIS except the inability to understand written and spoken Swedish for informed consent. Individuals were excluded from the S3WP study if they had: 1) previously

received health care for myocardial infarction, stroke, peripheral artery disease or diabetes, 2) presence of any clinically significant disease which, in the opinion of the investigator, may interfere with the results or the subject's ability to participate in the study, 3) any major surgical procedure or trauma within 4 weeks of the first study visit, or 4) medication for hypertension or hyperlipidemia. In addition to undergoing the examinations in SCAPIS, the 101 individuals in the S3WP study were screened at six visits over two years (Tebani et al., 2020). 12 individuals who received antibiotics during the course of the study were also excluded from gut microbiota analyses, and therefore we included samples from 75 of the 101 individuals who did not receive antibiotics and for whom we obtained fecal metagenomics samples from all four visits in the first year. Anthropometric and clinical chemistry parameters for the 75 individuals at the four visits are presented in Table S1A; gender was self-reported and no measurements were done to determine sex.

The study is approved by the Ethical Review Board of Gothenburg, Sweden Sweden (DNR 448-16, 407-15). All participants provided written informed consent. The study protocol conforms to the ethical guidelines of the 1975 Declaration of Helsinki.

METHOD DETAILS

Study visits in the S3WP study

Examinations were performed every third month +/- 2 weeks in the first year. All individuals fasted overnight (at least 8 hours) before the visits. Individuals underwent the same examinations at each visit, including measurements of body weight, waist and hip circumference, body fat using bioimpedance, and blood pressure as previously described (Tebani et al., 2020). A selection of questions from the initial SCAPIS questionnaire was repeated to note any changes in health and life-style factors between each visit such as infections, disease, medication and perceived health. Self-reported health issues during the study period were most commonly related to viral infections. Of the 75 individuals included in this study, 25 reported having symptoms of common cold and/or influenza, one reported having a *Campylobacter* infection, and two underwent any form of surgery (orthopedic hand surgery and resection of malignant melanoma).

Extraction of total fecal genomic DNA

Stool samples were collected at home and stored at room temperature for a maximum of 36 h before long-term storage at -80°C. Total genomic DNA was isolated from 100-120 mg of feces using a repeated bead beating method, as previously described (Wu et al., 2020). Briefly, fecal samples were placed in Lysing Matrix E tubes (MP Biomedicals) and extracted twice in lysis buffer (4% w/v SDS; 500 mmol/L NaCl; 50 mmol/L EDTA; 50 mmol/L Tris·HCl; pH 8) with bead beating at 5 m/s for 60 s in a FastPrep®-24 Instrument (MP Biomedicals). After each bead-beating cycle, samples were heated at 95°C for 5 min and then centrifuged at full speed for 5 min at 4°C. Supernatants from the two extractions were pooled and a 600 μL aliquot from each sample was purified using the QIAamp DNA Mini kit (QIAGEN) in the QIAcube (QIAGEN) instrument using the procedure for human DNA analysis. Samples were eluted in 200 μL of AE buffer (10 mmol/L Tris·Cl; 0.5 mmol/L EDTA; pH 9.0).

Metagenome sequencing of human fecal microbiota and analyses

1 μg of extracted total fecal genomic DNA was prepared for sequencing using Illumina TruSeq DNA PCR-Free sample prep kit and sequenced as 125 bp paired-end reads on an Illumina HiSeq 2500 sequencer.

Raw metagenome reads were analyzed using the MEDUSA platform (Karlsson et al., 2014), trimmed by FASTX with a quality threshold of 20 and minimum length of 35 bp, and filtered to remove human reads (NCBI version 37) using Bowtie2 (Langmead and Salzberg, 2012). High-quality reads devoid of human reads were aligned to genome and gene databases using Bowtie2 and with parameters -U - –no-sq –no-head -p 16 –very-sensitive. We obtained an average of 34.9±7.0 million high-quality reads per sample (10.5 billion reads in total for the 300 samples in our study). Mapping frequencies to genome and gene catalogs are indicated in Table S1H. The taxonomic information was obtained from a non-redundant catalog of 1639 genomes from NCBI Genbank database. Species level analyses were performed on a set of 384 species, from the 1629 genomes, that we could assign at least 10 reads on the genomes in at least 10% of the samples (i.e., 30 samples). Functional analysis was made using HUMAnN2 (Franzosa et al., 2018) and the MetaCyc database at the pathway level. We detected 466 annotated pathways in the 300 samples and performed analyses on the 331 pathways that were present in at least 20% of samples at an abundance higher than 0.0001%. Analyses on pathway composition were performed if more than 25% of the pathway abundance had a taxonomic classification in more than 25% of the samples. For targeted analysis of the terminal genes for butyrate synthesis (i.e. 5 genes *4hbt*, *but*, *atoA/D*, and *buk* representing 4 different pathways (Vital et al., 2014)), high-quality reads were aligned to a gene catalog (Karlsson et al., 2014) using Bowtie2 (Langmead and Salzberg, 2012). Hidden Markov models for the 5 genes (Vital et al., 2014) were used in the search using HMMER (hmmer.org, version 3.2.1) on gene sequences translated into amino acids (6-frames), and hits with full sequence score > 100 were retained. The abundances of pathways in each sample were determined by summing the counts of all genes in a samples that were annotated to a specific terminal gene.

16S rRNA gene sequencing of human fecal microbiota

The fecal microbiota was also profiled by sequencing the V4 region of the 16S rRNA gene on an Illumina MiSeq instrument (Illumina RTA v1.17.28; MCS v2.5) with 515F and 806R primers designed for dual indexing (Kozich et al., 2013) and the V2 MiSeq SBS sequencing reagents (paired-end, 250bp). 16S rRNA genes were amplified in duplicate reactions in volumes of 25 μL containing 1× Five Prime Hot Master Mix (5 PRIME GmbH), 200 nmol/L of each primer, 0.4 mg/mL BSA, 5% DMSO, and 20 ng of total fecal

genomic DNA. PCR was carried out under the following conditions: initial denaturation for 3 min at 94°C, followed by 25 cycles of denaturation for 45 s at 94°C, annealing for 60 s at 52°C, and elongation for 90 s at 72°C, and a final elongation step for 10 min at 72°C. Duplicates were combined, purified with the NucleoSpin Gel and PCR Clean-up kit (Macherey–Nagel), and quantified using the Quant-iT PicoGreen dsDNA kit (Invitrogen). Purified PCR products were diluted to 10 ng/μL and pooled in equal amounts. The pooled amplicons were purified again using Ampure magnetic purification beads (Agencourt) to remove short amplification products.

Illumina reads were merged using PEAR and quality filtered by removing all reads that had at least one base with a q-score lower than 20 (Zhang et al., 2014). Quality filtered reads were analyzed with the software package QIIME (version 1.9.1) (Caporaso et al., 2010b). Sequences were clustered into operational taxonomic units (OTUs) at a 97% identity threshold using an open-reference OTU picking approach with UCLUST (Edgar, 2010) against the Greengenes reference database, 13_8 release (DeSantis et al., 2006). Representative sequences for the OTUs were taxonomically assigned using the Greengenes taxonomy and the Ribosomal Database Project Classifier (Wang et al., 2007). Representative OTUs were aligned using PyNAST (Caporaso et al., 2010a) and used to build a phylogenetic tree with FastTree (Price et al., 2010), which was used to calculate α- and β-diversity of samples using Phylogenetic Diversity (Faith, 1994) and UniFrac (Lozupone and Knight, 2005). Chimeric sequences were identified with ChimeraSlayer (Haas et al., 2011) and excluded from all downstream analyses. Similarly, OTUs that could not be aligned with PyNAST, singletons and low abundant OTUs with a relative abundance <0.002% were also excluded. We obtained an average of 61,997 ± 17,134 sequences/sample (mean ± SD; range 16,663–144,133 sequences/sample); a total of 23,310,906 sequences and 1,466 OTUs were included in the analyses. To correct for differences in sequencing depth between samples, 16,600 sequences were randomly sub-sampled from each sample and included in the analyses for the estimation of α- and β-diversity.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were performed using the R environment and language (R Core Team, 2018).

Analysis of whole-genome metagenomic data was done at species level. Species abundances were filtered stringently based on both count abundances and prevalence: we included in the analyses species with more than 10 reads present in more than 20% of samples (to avoid bias in estimation of variability for low abundant taxa with abundances close to the limit of detection). Calculations were made on $\log_{10}$ transformed counts rarefied to 20 Mreads. Species relative abundances were calculated as counts for a species scaled to total count in the sample, with values summing up to 1.

The gut microbiota composition was analyzed with the Bray-Curtis dissimilarity index on counts rarefied to 20 Mreads, using vegdist function in the R-package *vegan* (Oksanen et al., 2015). Analysis of variance in composition was made on Bray-Curtis distances with permutational MANOVA using the function adonis2 in *vegan* with 9999 permutations. The intra-individual compositional variability was defined as the median Bray-Curtis dissimilarity calculated between samples for an individual (i.e., 6 dissimilarity values were calculated for the 4 samples obtained from each individual). The inter-individual compositional variability was defined as the median Bray-Curtis dissimilarity calculated for the 4 samples from an individual against all other samples.

To test the relationship between intra-individual gut microbiota compositional variation, gene richness, species abundances and variables such as taxonomy, functional potential, anthropometric variables and diet data sets, we used linear mixed-effects models with the *lme4* package (Bates et al., 2015). Median intra-individual Bray-Curtis dissimilarity, average gene richness (average of the 4 values for each individual), average *F. prausnitzii* abundance and fluctuation in *F. prausnitzii* (coefficient of variation of the 4 values for each individual) were used as fixed effects. P-values were FDR-adjusted using the p-value distribution (Storey et al., 2015) and q-values are indicated in the results. Normal ranges for individual species were defined as the central 95% of the population (between the 2.5^{th} and 97.5^{th} percentiles) for all 300 samples, using the *quantile* function with type 7 in R.

Mean and variance for each species were modelled, using individuals as random variable, on $\log_{10}$-transformed counts from the 4 visits for the 75 individuals using mixed-effect models without fixed effect. The total variance was partitioned into the inter-individual variance (σ^2_{BS}) and intra-individual variance (σ^2_{WS}) and ICC was calculated as $\sigma^2_{BS}/(\sigma^2_{BS+}\ \sigma^2_{WS})$.

Permutated Spearman's correlations were calculated using delta values between visits for species relative abundances and gene richness. Absolute values of delta relative abundance between visits for species were used in correlations for compositional dissimilarities. 9999 permutations, within individuals, were made for statistical evaluation.

Co-occurrence networks were built using the average (of the 4 values for each individual) correlations values between species abundances (Poyet et al., 2019) using the SparCC (Friedman and Alm, 2012) implemented in the SpiecEasi package for R (Kurtz et al., 2015) using default settings. Correlations with a coefficient larger than 0.5 or smaller than -0.5 were visualized using the igraph package (Csardi and Nepusz, 2006).

Community clustering was determined using Dirichlet multinomial mixtures (DMM) clustering (Holmes et al., 2012) for all 300 samples at genus level. The amount of constrained variation in microbiota composition was determined using capscale in *vegan* using Bray-Curtis dissimilarity and was visualized using principal coordinate analysis (PCoA). Differences in parameters between the enterotypes were tested using the Kruskal-Wallis test.