OPEN ACCESS

Original research

Long-term instability of the intestinal microbiome is associated with metabolic liver disease, low microbiota diversity, diabetes mellitus and impaired exocrine pancreatic function

Fabian Frost [1], Tim Kacprowski,[2] Malte Rühlemann [3], Maik Pietzner,[4] Corinna Bang [3], Andre Franke [3], Matthias Nauck,[4,5] Uwe Völker,[6] Henry Völzke,[7] Marcus Dörr,[5,8] Jan Baumbach,[2] Matthias Sendler,[1] Christian Schulz,[9] Julia Mayerle [1,9], Frank U Weiss,[1] Georg Homuth,[6] Markus M Lerch [1]

Check for updates

ABSTRACT

Objective The intestinal microbiome affects the prevalence and pathophysiology of a variety of diseases ranging from inflammation to cancer. A reduced taxonomic or functional diversity of the microbiome was often observed in association with poorer health outcomes or disease in general. Conversely, factors or manifest diseases that determine the long-term stability or instability of the microbiome are largely unknown. We aimed to identify disease-relevant phenotypes associated with faecal microbiota (in-)stability.

Design A total of 2564 paired faecal samples from 1282 participants of the population-based Study of Health in Pomerania (SHIP) were collected at a 5-year (median) interval and microbiota profiles determined by *16S rRNA* gene sequencing. The changes in faecal microbiota over time were associated with highly standardised and comprehensive phenotypic data to determine factors related to microbiota (in-)stability.

Results The overall microbiome landscape remained remarkably stable over time. The greatest microbiome instability was associated with factors contributing to metabolic syndrome such as fatty liver disease and diabetes mellitus. These, in turn, were associated with an increase in facultative pathogens such as *Enterobacteriaceae* or *Escherichia/Shigella*. Greatest stability of the microbiome was determined by higher initial alpha diversity, female sex, high household income and preserved exocrine pancreatic function. Participants who newly developed fatty liver disease or diabetes during the 5-year follow-up already displayed significant microbiota changes at study entry when the diseases were absent.

Conclusion This study identifies distinct components of metabolic liver disease to be associated with instability of the intestinal microbiome, increased abundance of facultative pathogens and thus greater susceptibility toward dysbiosis-associated diseases.

Significance of this study

What is already known on this subject?
► The gut microbiome represents a complex ecosystem exerting vital functions and its disturbance can contribute to the pathogenesis of many diseases—particularly when microbiome diversity is reduced.
► Little, however, is known about which host phenotypes and diseases affect the long-term stability of the gut microbiome or render it instable.

What are the new findings?
► At population level, the overall faecal microbiome community structure is remarkably stable over time.
► Individuals suffering from metabolic liver disease, diabetes mellitus or related phenotypes show the vastest long-term microbiome perturbations.
► Factors associated with the greatest long-term stability of the faecal microbiota composition are high initial microbial (alpha) diversity, female sex, high household income and preserved exocrine pancreatic function.
► Microbiota instability in general, and specifically fatty liver disease and diabetes mellitus are linked to a long-term increase in opportunistic pathogens like *Enterobacteriaceae* or *Escherichia/Shigella*.
► A specific gut microbiota signature is associated with an increased susceptibility toward incident metabolic liver disease.

INTRODUCTION

The human gut constitutes a habitat for the complex ecosystem which encompasses trillions of micro-organisms, together known as the intestinal microbiome. Bacterial cells represent the largest fraction of this biotope even outnumbering their human host by cell number.[1] Jointly the gut microbiota carry out important metabolic functions,[2] which are critical for the maintenance of human health.[3] At the same time dysbiosis, a state of structural and functional microbiota disruption, can contribute to a variety of

Significance of this study

How might it impact on clinical practice in the foreseeable future?
► Individuals suffering from metabolic liver disease or metabolic syndrome may exhibit greater susceptibility towards dysbiosis-associated disorders.
► Whether attempts to maintain or restore the gut microbiome community structure can positively affect the onset and course of metabolic liver disease needs to be investigated in interventional trials.

disorders including metabolic liver disease.[4 5] The gut microbiota structure is not static over time but subject to continuous adaptive changes. The high interpersonal gut microbiota variation at a very young age decreases substantially until adulthood when the gut microbiome becomes more stable.[6] But even afterwards, marked alterations as a response to environmental stimuli, for example, changes in diet, are possible but can still be largely reversible if the trigger is discontinued.[7] Other environmental changes such as permanent shifts in geographic location, however, may cause enduring gut microbiota alterations.[8] Ultimately, over time each individual faces numerous life events which will cumulatively shape the individual gut microbiota profiles. Previous investigations which tried to determine the specific changes occurring in the gut microbiota during adulthood mostly relied on cross-sectional data.[6 9] This approach can be problematic as people of different ages have often been subjected to very different environmental conditions in the early years of their lives, which is a period of particular sensibility for the initial gut microbiota constitution.[10] Moreover, cross-sectional analyses cannot determine whether and how the presence of an overt disease subsequently affects the microbiota composition and stability. Intraindividual comparisons using paired long-term follow-up data are therefore necessary to determine the contribution of different phenotypic factors to the gut microbiome changes over time. We therefore investigated intestinal microbiota profiles by *16S rRNA* gene sequencing in 2564 paired faecal samples from 1282 participants of the Study of Health in Pomerania (SHIP), which were collected at two time points with a 5-year interval (SHIP-2 and SHIP-3, respectively) in this longitudinal cohort study (figure 1).

METHODS

Study participants
SHIP is a longitudinal cohort study located in Northeast Germany.[11] It was designed to determine the incidence and prevalence of subclinical disorders, clinical disease and their risk factors.[12 13] The population-based sample was drawn in 1996 from a population of 213 057 using a two-stage stratified cluster approach including only individuals aged 20–79 years. The net sample included 6265 subjects of which 4308 participated in the study. Follow-up investigations are scheduled in 5-year intervals and have already been performed from 2002 to 2006 (SHIP-1, n=3300), 2008–2012 (SHIP-2, n=2333) and 2014–2016 (SHIP-3, n=1718). Faecal samples for microbiota determination were collected at SHIP-2 and SHIP-3. Paired faecal microbiota data at both study time points were available for 1366 individuals. Of these, 84 datasets were removed due to intake of antibiotics at the time of sample collection, a history of gastrointestinal cancer, liver cirrhosis or missing data on a history of
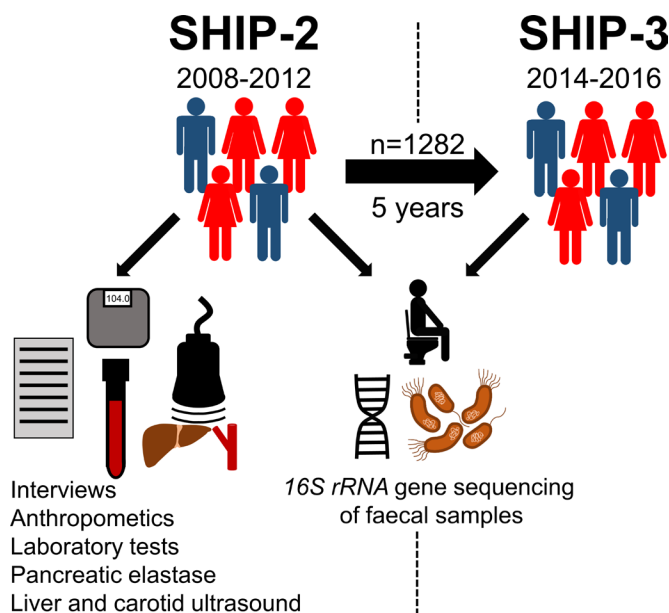

Figure 1 Scheme of study design. Paired microbiota profiles from Study of Health in Pomerania (SHIP)-2 and SHIP-3 participants (n=1282 at each time point) and corresponding SHIP-2 phenotype data were included into the analysis.

liver cirrhosis, leaving 1282 datasets for analysis. All participants provided written informed consent and the study was approved by the local ethics committee.

16S rRNA gene sequencing
Sequencing was performed as described before in detail.[14 15] In brief, faecal samples were collected by the study participants at home, stored in a tube containing stabilising DNA buffer and then transported to our laboratory by the participants or courier. After DNA from faecal samples was isolated (PSP Spin Stool DNA Kit; Stratec Biomedical AG, Birkenfeld, Germany), it was stored at −20°C until analysis by *16S rRNA* gene sequencing of the V1–V2 region on a MiSeq platform (Illumina, San Diego, California, USA).

Assignment of taxonomic annotation and predicted metagenomics
More details on taxonomy assignment and predicted metagenomics are given in the online supplemental methods. In brief, the open-source software package DADA2[16] (V.1.10) was used for amplicon data processing which enables a single-nucleotide resolution of amplicons. The functional genomic potential of the faecal microbiota were predicted with PICRUSt2.[17] All samples were normalised to 10 000 *16S rRNA* gene read counts for analysis.

Metabolomic measurements
In a subsample of 911 SHIP-2 individuals plasma metabolites were determined using the AbsoluteIDQ p180 Kit (BIOCRATES LifeSciences AG, Innsbruck, Austria) on an AB SCIEX 5500 QTrap mass spectrometer (AB SCIEX, Darmstadt, Germany) with electrospray ionisation combined with a HPLC system (Agilent 1260 Infinity Binary LC, Santa Clara, California, USA). Following quality control and data processing, 177 metabolites were available for statistical analyses.
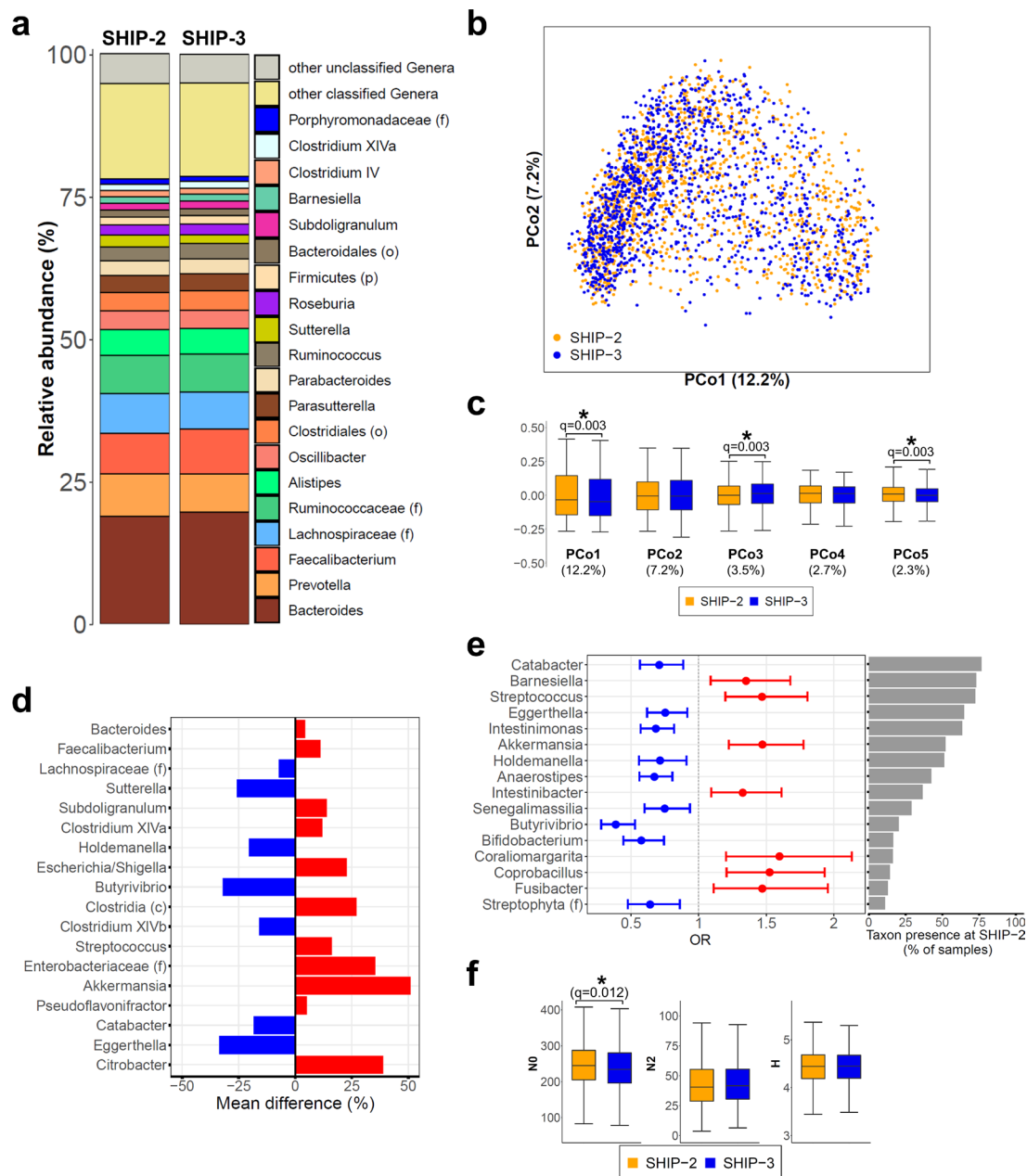
Figure 2 Stability of the faecal microbiome from Study of Health in Pomerania (SHIP)-2 to SHIP-3. (A) Stacked bar plots show the average faecal microbiota composition at SHIP-2 and SHIP-3. (B) Principal coordinate analysis (PCoA) of 2564 faecal microbiota samples obtained from 1282 individuals at SHIP-2 (orange dots) and SHIP-3 (blue dots). (C) Boxplots display the PCo1–PCo5 scores comparing SHIP-2 (orange boxes) to SHIP-3 (blue boxes) samples. Numbers in brackets below denote the percentage of variation explained by the respective PCo. The overall faecal microbiome landscape was quite stable from SHIP-2 to SHIP-3 with small but significant shifts along PCo1, PCo3 and PCo5. (D) Shown are the percentage changes in abundance of the 18 taxa with a significant increase (red) or decrease (blue) in abundance over time. All taxa are ordered according their mean abundance at SHIP-2 from top (most abundant) to bottom (least abundant). (E) Analysis of the taxon presence distribution from SHIP-2 to SHIP-3. Shown are the ORs and their respective 95% CIs at SHIP-3 compared with SHIP-2 as baseline of all 16 taxa with significant changes in their presence profiles. An OR over 1 indicates an increased likelihood for the respective taxon to be present in a SHIP-3 sample if it was not present in the same subject at SHIP-2, whereas an OR below 1 implies the loss of the respective taxon in SHIP-3 samples compared with SHIP-2. All taxa are ordered according to the percentage of samples with presence of the respective taxon at SHIP-2 from top (high presence) to bottom (low presence). (F) Comparison of alpha diversity between SHIP-2 and SHIP-3. Species richness (N0) was slightly reduced at follow-up, whereas no change was observed in Simpson diversity number (N2) or Shannon diversity index (H). Whiskers are drawn up to 1.5 times the IQR (outliers not shown). *Indicates significant difference. c, class; f, family; o, order; p, phylum.

Phenotype data

Determination of laboratory parameters (alanine aminotransferase (ALT), creatinine, high-density lipoprotein (HDL), low-density lipoprotein (LDL), N-terminal prohormone of brain natriuretic peptide (NT-proBNP), thyroid-stimulating hormone (TSH)) was performed on a Dimension VISTA platform (Siemens Healthcare Diagnostics, Eschborn, Germany). Glycated haemoglobin (HbA1c) concentrations were determined using a

high-performance liquid chromatography (Bio-Rad Diamat, Munich, Germany). Estimation of glomerular filtration rate (eGFR) was done using the CKD-EPI equation.[18] NT-proBNP measurements were treated as missing in case of eGFR ≤30 mL/min due to possible accumulation. The trait diabetes mellitus was assigned to all individuals with a positive history of diabetes mellitus in combination with a history of current treatment (diet, medication or insulin), a random blood glucose ≥11.1 mmol/L or an HbA1c ≥6.5%. Alcohol consumption was estimated in grams alcohol per day, by counting the number of alcoholic beverages that had been consumed on average per day over the last 30 days and calculation of its average alcoholic content. Smoking was quantified as on average smoked cigarettes per day. Fatty liver disease (FLD) was determined based on ultrasound findings in case of hyperechoic liver tissue (in the absence of a history of liver cirrhosis) using a high-resolution device (Vivid i, GE Healthcare, Chicago, Illinois, USA) in B mode in combination with serum ALT levels belonging to the upper 25% of the investigated population. The trait arterial hypertension was assigned in case of antihypertensive medication or a systolic blood pressure ≥140 mm Hg or a diastolic blood pressure ≥90 mm Hg in the third of three consecutive measurements. Carotid intima–media thickness was measured bilaterally (Vivid i) and thereafter the average of both sides calculated. Manifest atherosclerotic disease (as opposed to carotid intima–media thickness) was assigned if the participants had a history of myocardial infarction and/or stroke. Hypothyroidism was assumed when TSH was ≥4 mU/L or in case of thyroid hormone replacement therapy. A food frequency score was used to estimate the balance of diet, with higher food frequency scores suggesting a more balanced diet as described elsewhere.[19 20] For calculation of this score, ordinal consumption data from 15 food categories (meat, sausage, fish, boiled potatoes, pasta, rice, raw vegetables, boiled vegetables, fruits, whole grain/black/crisp bread, wheat oats/cornflakes, eggs, cake/cookies, sweets and savoury snacks) was evaluated. The net household income was evaluated in classes ('1': <€500, '2': €500–€900, '3': €900–€1300, '4': €1300–€1800, '5': €1800–€2300, '6': €2300–€2800, '7': €2800–€3300, '8': €3300–€3800, '9': >€3800). Exocrine pancreatic function was determined using a pancreatic elastase ELISA (BIOSERV Diagnostics, Germany) based on faecal samples as described before.[14]

Data analysis

All statistical analyses were performed using 'R' (v.3.6.3).[21] Alpha diversity scores were determined with the 'diversity' function (vegan).[22] The Bray-Curtis (BC) dissimilarity, Euclidean and Jaccard distance were calculated using the function 'vegdist' (vegan). Principal coordinate analysis (PCoA) was carried out by the 'cmdscale' function (vegan). The emergence of negative eigenvalues during PCoA was avoided by square root transformation of the BC dissimilarity prior to the ordination. To test for significant differences of the PCo1–PCo5 scores between SHIP-2 and SHIP-3, paired two-sided Wilcoxon-signed rank tests were performed. For permutational analysis of variance, the function 'adonis' (vegan; 1000 permutations) was used. Comparison of the relative abundance data between SHIP-2 and SHIP-3 of all taxa that were present in at least 10% of all SHIP-2 samples and with a mean abundance of at least 0.1% was done using the paired two-sided Wilcoxon-signed rank tests. Comparison of the taxon presence–absence profiles between SHIP-2 and SHIP-3 of all taxa that were present in at least 10% but less than 80% of all samples was performed using McNemar's test ('mcnemar.exact' function, exact2×2). Differences in genomic pathways between

SHIP-2 and SHIP-3 were evaluated using the paired two-sided Wilcoxon-signed rank tests. Only pathways with a presence of at least 25% of samples at SHIP-2 were considered for this analysis. Alpha diversity scores of SHIP-2 and SHIP-3 samples were compared using the paired two-sided Wilcoxon-signed rank test. Assessment of association between phenotypic factors at SHIP-2 and faecal microbiota changes toward SHIP-3 (microbial (in-)stability) expressed as log-transformed BC dissimilarity, Euclidean or Jaccard distance was performed as follows: metric variables were cleaned from outliers which were more than 3 SD away from the mean. Metric variables without zeros (eg, HbA1c) were log-transformed, whereas square root transformation was performed in case of present zeros (cigarettes per day or alcohol consumption). Afterwards, Pearson, Spearman or point biserial correlations were calculated for metric (age, alcohol, body mass index (BMI), carotid intima–media thickness, HbA1c, LDL/HDL ratio, eGFR, pancreatic elastase, number of household members, NT-proBNP, species richness (N0), smoking, TSH), ordinal (food frequency score, net household income) or nominal data (arterial hypertension, diabetes mellitus, dyslipidaemia, FLD, female sex, hypothyroidism, manifest atherosclerotic disease, proton pump inhibitor usage), respectively ('cor.test' function, stats). Afterwards, the significance of those factors identified to correlate with microbiota (in-)stability was confirmed by performing a linear regression model including possible confounding factors as stated in the online supplemental table S4. The associations of incident diseases at SHIP-3 with SHIP-2 and SHIP-3 microbiota profiles were determined by performing permutational analysis of variance (vegan function 'adonis'; 1000 permutations). For linear regression analysis, microbiota data were cleared from zeros (treated as missing), log-transformed and standardised. Metabolome data were also log-transformed and standardised before linear regression analysis. Included covariates for regression models are listed in the online supplemental tables S4–S11 of the respective analyses. All p-values derived from comparisons of PCo scores, phenotype–microbiota associations, taxon comparisons of relative abundance, presence–absence, genomic pathway or alpha diversity data were corrected for multiple testing by the method of Benjamini-Hochberg ('p.adjust' function, stats) and thereafter called q-values. P-values or q-values <0.05 were considered significant. All p-values and q-values were rounded to three significant digits.

Data availability

All microbiome and phenotype data were obtained from the SHIP data management unit and can be applied for online through a data access application form (https://www.fvcm.med.uni-greifswald.de/dd_service/data_use_intro.php).

Patient and public involvement statement

SHIP is a regionally confined epidemiological study that regularly informs participants about current investigations and recent findings by newsletter, YouTube videos (eg, https://www.youtube.com/watch?v=t7C8dGNJ_rk) and its website (https://www2.medizin.uni-greifswald.de/cm/fv/ship/). Participants serve on the advisory board of the SHIP study and SHIP organises public hearings for question and answer sessions and for input from the public regarding its research focus.

RESULTS
Overall faecal microbiota community profile

The overall faecal microbiota community profile was remarkably stable over time with *Bacteroides*, *Prevotella* and *Faecalibacterium*
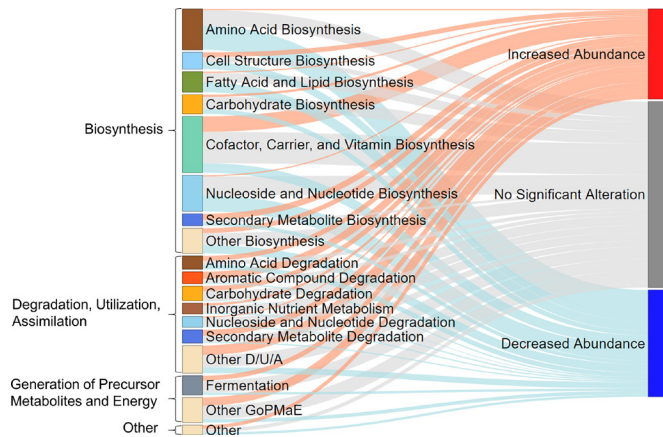
Figure 3 Sankey plot displays the proportions of predicted metagenomic pathways with significantly (q<0.05) increased or decreased abundance or no significant alteration. D/U/A, degradation, utilisation, assimilation; GoPMaE, generation of precursor metabolites and energy.

remaining the three most abundant taxa at follow-up (figure 2A). To further investigate the temporal stability of the faecal microbiota, an analysis of beta diversity, which provides information how microbiota communities or samples differ from each other, was carried out using BC dissimilarity. Subsequent PCoA based on BC including all paired 2564 baseline and follow-up samples confirmed a largely stable faecal microbiota composition between SHIP-2 and SHIP-3 (figure 2B). However, detailed analysis of the PCo scores revealed significant structural microbiota changes (figure 2C) indicating specific microbiota shifts in some individuals.

Determination of changes in taxon abundance and predicted metagenomics from baseline to follow-up

A total of 18 taxa were identified with significant abundance changes at follow-up compared with baseline (figure 2D and online supplemental table S1). A comparison between baseline and follow-up samples based on presence–absence data revealed 16 taxa with significantly altered presence profiles (figure 2E and online supplemental table S2). No major changes were found in alpha diversity, which describe the taxon variation within a sample, although median N0 was slightly lower at follow-up (q=0.012) compared with baseline (figure 2F). Corresponding to the changes in several taxa, alterations in 164 predicted genomic microbial pathways were found (online supplemental table S3, figures S1 and S2). These pathways belonged to the MetaCyc[23] superclasses biosynthesis (n=100), degradation/utilisation/assimilation (n=43), generation of precursor metabolites and energy (n=16) and others (n=5) (figure 3). Several of the metabolic pathways for short-chain fatty acid (SCFA) production were either increased or decreased in abundance, similarly to the taxonomic changes in SCFA producers, yet the combined predicted microbial genetic potential for SCFA production, however, did not change. Analysis of taxon abundances revealed several Gram-negative bacteria such as *Enterobacteriaceae*, *Escherichia/Shigella* and *Citrobacter* to be increased at follow-up (figure 2D). Correspondingly, the genetic microbial potential for the production of lipopolysaccharides (LPS) was increased at SHIP-3 compared with SHIP-2.

Table 1 Distribution of phenotype variables at SHIP-2

Continuous variables	Mean (±SD)	Median (first–third quartile)	Missing data (%)
Age (years)	55.8 (±12.3)	56.0 (46.0–65.0)	0.0
Alcohol (g/day)	9.1 (±12.7)	4.3 (1.2–11.5)	2.2
Balanced diet (Food frequency score)	14.5 (±3.4)	15.0 (12.0–17.0)	0.0
BMI (kg/m²)	28.0 (±4.8)	27.4 (24.6–30.7)	0.1
Carotid intima–media thickness (mm)	0.6 (±0.1)	0.6 (0.5–0.7)	1.2
Exocrine pancreatic function (pancreatic elastase, μg/g stool)	460.3 (±166.8)	481.0 (346.0–585.0)	0.5
Glycaemic control (HbA1c, %)	5.4 (±0.8)	5.3 (5.0–5.7)	0.1
Cardiac function (NT-proBNP, pg/mL)	112.6 (±174.4)	66.5 (36.0–120.8)	3.1
Household net income (class)	4.9 (±2.0)	5.0 (4.0–6.0)	1.5
LDL/HDL ratio	2.5 (±0.9)	2.3 (1.8–3.1)	0.2
No of household members	2.2 (±0.9)	2.0 (2.0–3.0)	0.0
Renal function (eGFR, mL/min)	87.6 (±17.2)	88.5 (76.3–99.9)	0.1
Smoking (cigarettes/day)	2.1 (±5.5)	0.0 (0.0–0.0)	0.0
Species richness (N0)	248.3 (±63.4)	245.0 (205.0–287.0)	0.0
Thyroid function (TSH, mU/L)	1.2 (±1.0)	1.0 (0.7–1.5)	0.3

Binary variables	Positive reports (%)	Missing data (%)
Arterial hypertension	48.8	0.1
Diabetes mellitus	10.6	0.1
Dyslipidaemia	34.8	0.2
Fatty liver disease	16.8	0.8
Female sex	53.9	0.0
Hypothyroidism	12.7	0.3
Manifest atherosclerotic disease	2.3	0.0
Proton pump inhibitor usage	10.5	0.0

All variables are in alphabetical order. All values are rounded to one decimal place.
BMI, body mass index; eGFR, estimated glomerular filtration rate; HbA1c, glycated haemoglobin; HDL, high-density lipoprotein; LDL, low-density lipoprotein; ; NT-proBNP, N-terminal prohormone of brain natriuretic peptide; SHIP, Study of Health in Pomerania; TSH, thyroid-stimulating hormone.

Determination of phenotypic factors associated with stability or instability of the faecal microbiome

Several mostly cross-sectional studies have been conducted to identify the factors that contribute to the gut microbiota variation showing associations with age, sex, exocrine pancreatic function, obesity, diabetes mellitus, FLD, diet or medication among other factors.[7 14 24–27] We therefore investigated whether these and other phenotypic factors (table 1) were also associated either positively or negatively to longitudinal faecal microbiota changes from baseline to follow-up (expressed as beta diversity). As a result, numerous host characteristics were found to be associated with either increased or decreased faecal microbiota stability (figure 4). A positive correlation with changes in the faecal microbiome (BC) over time was shown for FLD (q=0.006), diabetes mellitus (q=0.007) and the level of glycaemic control (HbA1c, q=0.014). An increase in BMI exhibited a smaller positive but non-significant (q=0.067) association. Several other phenotypic traits were predictive of faecal microbiota stability (negative association to BC), among them high initial microbial diversity (N0 at SHIP-2, q<0.001), female
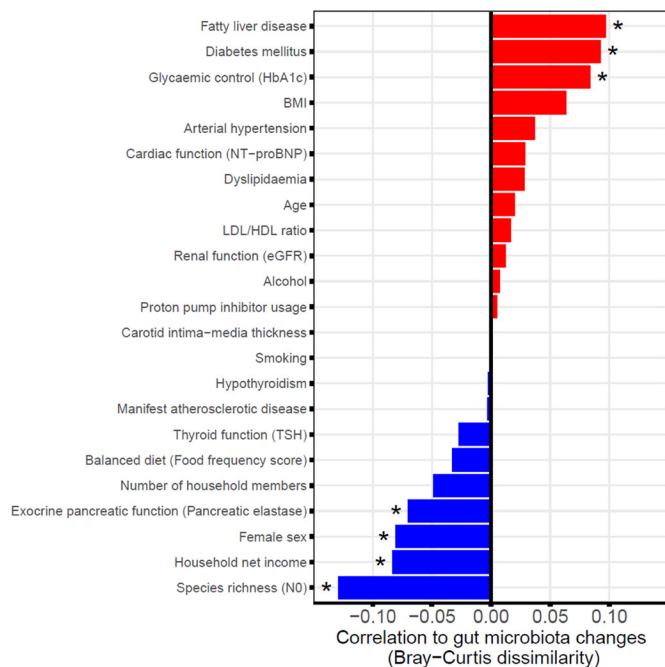
Figure 4 Contribution of phenotypic factors to faecal microbiome changes. Shown are the correlations between different phenotypic factors at SHIP-2 and global changes in the faecal microbiome from SHIP-2 to SHIP-3. Large positive associations indicate greater variation or instability in the faecal microbiome if the respective factor is distinctly elevated (continuous variables) or positive (binary variables). On the other hand, large negative associations suggest stability of the faecal microbiome. *Indicates significant associations (q<0.05). BMI, body mass index; eGFR, estimated glomerular filtration rate; HbA1c, glycated haemoglobin; HDL, high-density lipoprotein; LDL, low-density lipoprotein; NT-proBNP, N-terminal prohormone of brain natriuretic peptide; SHIP, Study of Health in Pomerania; TSH, thyroid-stimulating hormone.

sex (q=0.015), high net household income (q=0.014) and preserved exocrine pancreatic function (q=0.039). Additional regression analyses with adjustment for possible confounding factors (online supplemental table S4A), including the oral antidiabetic medications metformin[28] and dipeptidyl peptidase-4 (DPP-4) inhibitors[29] in case of the traits diabetes mellitus and glycaemic control confirmed the significance of the aforementioned associations. Using Jaccard distance (online supplemental figure S3A and table S4B) as measurement of beta diversity, which only includes presence/absence information of the bacteria, additionally revealed PPI usage to be positively associated with larger gut microbiota changes whereas a balanced diet was linked to gut microbiota stability over time. Using the Euclidean distance instead did not reveal further additional findings (online supplemental figure S3B and table S4C).

Contribution of microbial instability, FLD and diabetes mellitus to levels of facultative pathogenic bacteria

Taxon abundance comparisons between baseline and follow-up exhibited an increase of several facultative pathogenic organisms (*Enterobacteriaceae*, *Escherichia/Shigella* and *Citrobacter*) over time (p<0.001) (online supplemental figure S4 and table S5). Using a linear regression model, not only microbial instability (BC, p=0.027) but also FLD (p=0.038) and diabetes mellitus (p=0.007) were associated

with the increase in abundance of facultative pathogens toward SHIP-3. Correspondingly, microbial instability (p=0.013) and diabetes mellitus (p=0.021) were also associated with an increase in the predicted microbial potential for LPS biosynthesis toward SHIP-3.

Incident metabolic and cardiovascular disease at SHIP-3 and its association to faecal microbiota and plasma metabolites

A significant proportion of participants who did not show signs of diabetes mellitus, FLD, dyslipidaemia, hypertension, manifest atherosclerosis or hypothyroidism were classified as having acquired the respective condition 5 years later in SHIP-3 (figure 5A and online supplemental table S6). We investigated whether the microbiota profiles of those individuals who had newly developed any of these diseases in SHIP-3 differed from those who remained disease free at both time points. Individuals newly classified with FLD or diabetes mellitus at SHIP-3 exhibited a significantly different gut microbiome composition compared with participants without the respective disorder (p<0.001 and p=0.008, respectively). This association could already be found at SHIP-2 before FLD or diabetes mellitus had become apparent (p=0.009 and p=0.034). Incident hypertension (p=0.003) also exhibited a significant link to the gut microbiome at SHIP-3, but not in SHIP-2. No significant link to the faecal microbiome was found for incident cases of manifest atherosclerosis, dyslipidaemia or hypothyroidism. This indicates that participants who newly developed diabetes mellitus or FLD during the 5-year follow-up already displayed significant microbiota composition changes at study entry when the diseases were still absent.

To further explore these changes in individuals with incident FLD or diabetes mellitus, we performed linear regression analysis adjusting for different possible confounding factors (online supplemental tables S7–11). Individuals with incident FLD at SHIP-3 already presented with changes in *Clostridium XIVa*, *Clostridium XIVb*, *Collinsella* and *Oscillibacter* as well as 15 plasma metabolite levels prior to their diagnosis at SHIP-2 (figure 5B,C and online supplemental tables S7 and 8). *Clostridium XIVa* did not only associate with increased plasma levels of various phosphatidylcholines but also exhibited a strong association with an increase of gut microbial pathways for fatty acid biosynthesis which was independent of the individual's diet (food frequency score) and BMI (online supplemental table S9). The gut microbiota and plasma metabolome alterations at SHIP-2 in individuals with incident diabetes mellitus at SHIP-3 were of lesser extent (figure 5D,E and online supplemental tables S10 and S11).

DISCUSSION

In the present study, we performed a long-term intestinal microbiota follow-up of 1282 individuals to identify the phenotype factors that (de-)stabilise the microbiome. Overall, the global faecal microbiota profile of the whole population showed a largely stable composition over time. Levels of alpha diversity remained stable except for a small reduction in N0. However, levels of facultative pathogenic taxa such as *Enterobacteriaceae*, *Escherichia/Shigella* and *Citrobacter* were consistently increased at follow-up after 5 years. This finding was pronounced in subjects carrying a phenotype associated with less stable microbiota composition. Members of the family of *Enterobacteriaceae*, especially *Escherichia coli*, represent the most common pathogens found in either blood or bile from individuals suffering from acute cholecystitis or cholangitis,[30 31] abdominal
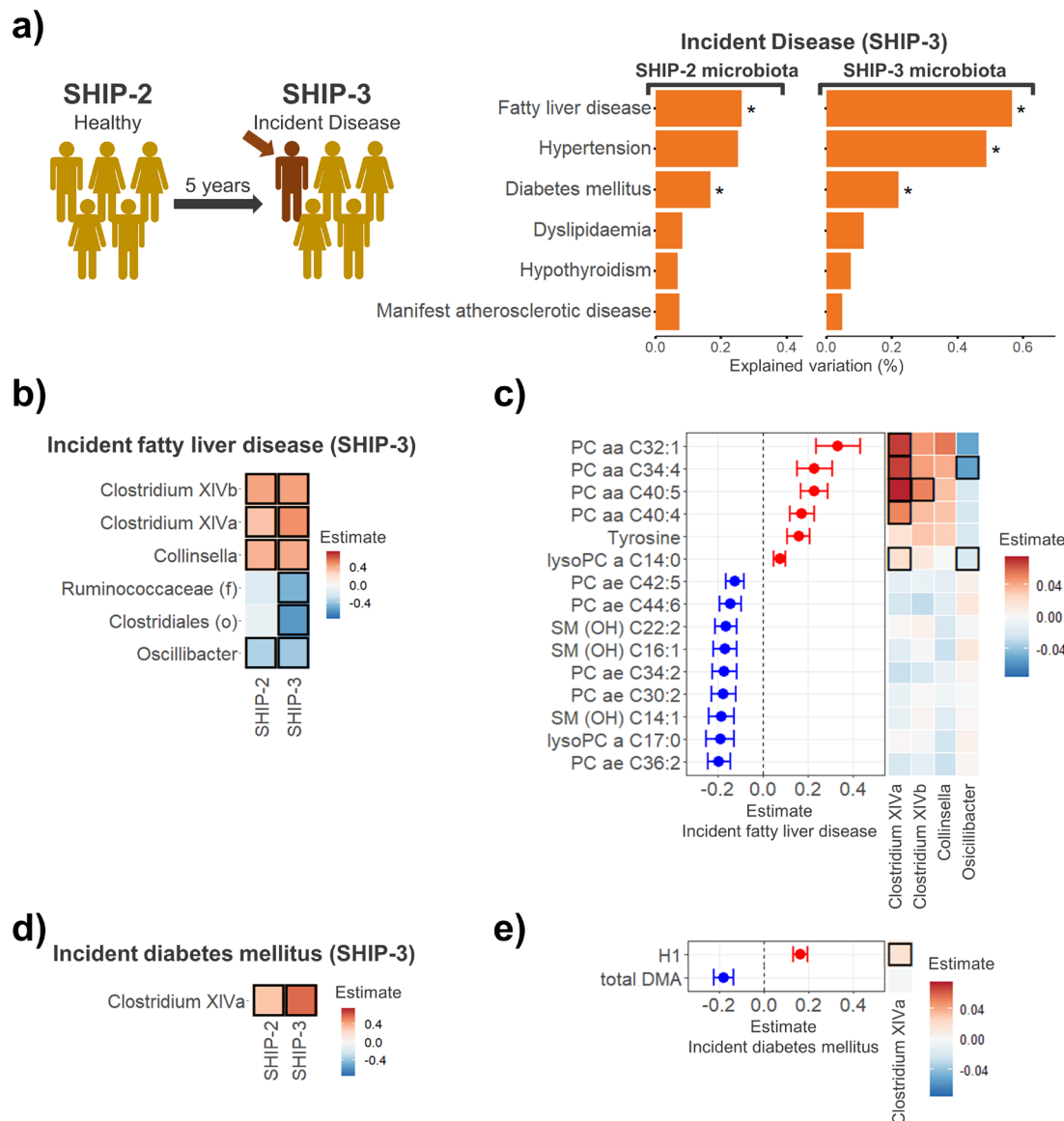
Figure 5 Associations between incident metabolic and cardiovascular diseases at SHIP-3 and faecal microbiota changes in the affected individuals at SHIP-2 (before the respective disease was present). (A) Individuals with incident fatty liver disease and diabetes mellitus at SHIP-3 already displayed gut microbiota changes at SHIP-2 prior to their diagnosis. (B) Associations of incident fatty liver disease with microbial taxa at SHIP-2 and SHIP-3. (C) Significant metabolite associations of incident fatty liver disease cases at SHIP-2 (left). The heatmap (right) indicates the association of the different disease-associated genera with the respective metabolites. (D) Associations of incident diabetes mellitus at SHIP-3 with microbial taxa at SHIP-2 and SHIP-3. (E) Significant metabolite associations of incident diabetes mellitus cases at SHIP-2 (left). The heatmap (right) indicates the association of the different disease-associated genera with the respective metabolites. Significant results are highlighted by a black frame. Unclassified taxa at genus level: (f): family, (o): order. DMA, dimethylarginine; H1, hexose; PC, phosphatidylcholine; SHIP, Study of Health in Pomerania; SM (OH), hydroxylated sphingomyelin.

abscesses (eg, liver abscesses)[32] or urinary tract infections.[33] Similarly *Citrobacter*, another member of *Enterobacteriaceae*, has been identified in urinary tract infections or intra-abdominal abscesses.[34] The order *Enterobacterales*, which comprises the aforementioned facultative pathogens, has also been linked to increased intestinal permeability in humans.[35] Therefore, greater microbiota instability with increased abundance of these potential pathogens may have harmful consequences for its human hosts by increasing the risk for localised or systemic infections. Moreover, at follow-up the predicted microbial pathways for LPS biosynthesis were increased by 27%. This increase was again pronounced in subjects with less stable faecal microbiota composition, for example, diabetics. LPS triggers the production

of proinflammatory cytokines and can contribute to the development of FLD and diabetes mellitus.[36–38] Hence, the microbial instability may predispose to further increases in LPS secretion activity which in turn can aggravate the metabolic disorders in a vicious circle.

A further, probably disadvantageous, faecal microbiota alteration was a reduction in the amount of *Bifidobacteria* carriers and the predicted genetic potential for bifidobacterial lactate production which could not be attributed to other phenotypic factors. *Bifidobacteria* have been reported to stabilise the gut barrier by reducing its permeability for inflammatory bacterial molecules such as LPS and to alleviate LPS-associated inflammation in the liver.[39–41]

The individual extent of the faecal microbiota variation between SHIP-2 and SHIP-3 was variable. Disorders contributing to metabolic syndrome, metabolic liver disease and diabetes mellitus were associated with faecal microbiota instability. In our study, the association of diabetes mellitus with faecal microbiota variation could not be explained by oral antidiabetic medication such as metformin or DPP-4 inhibitors, both of which have previously been shown to be independently associated with composition changes of the gut microbiome.[28 29] When analysing changes in presence/absence patterns, we also found PPI usage to be linked to gut microbiota instability over time, which supports the finding of gut microbiota changes in PPI usage from a previous cross-sectional study.[42] Of note, other conditions such as atherosclerosis, cardiac function, thyroid dysfunction or renal impairment had no significant impact. Participants' older age was not associated with faecal microbiota instability. This suggests that as long as metabolic disease is absent, even at higher age, the gut microbiome may remain largely stable and resilient to perturbations.

On the other hand, several factors which were related to greater microbiota stability could be identified. The strongest negative association was found for alpha diversity at study entry, implying a microbiome stabilising effect. Diverse communities are less prone to the introduction of foreign species because of a greater coverage of ecological niches.[43] A prominent example of how low microbial diversity can negatively affect the gut microbial integrity is *Clostridium difficile*, for which in vitro studies have shown an increased growth in dysbiotic microbiota with low diversity.[44] In humans, this situation typically occurs after antibiotic administration, a risk factor for *Clostridium difficile*-associated colitis. A further protective factor of the faecal microbiome stability was preserved exocrine pancreatic function. Exocrine pancreatic function has already been shown to be the most important host factor regulating the faecal microbiota composition in a given population,[14] which now includes a prominent role in determining long-term stability of the microbiome. This relationship may be explained by changes in the intestinal chyme and nutrient supply, due to maldigestion when pancreatic enzyme activity is reduced or due to impaired secretion of pancreatic antimicrobial peptides.[45] The increasing faecal microbiota stability seen in individuals with higher socioeconomic status (net income) may be mediated by a reduced prevalence of disease in general, an increased utilisation of health preventive measures and/ or a healthier lifestyle.

Interestingly, individuals who newly developed FLD (and to a lesser degree also diabetes mellitus) at SHIP-3 already displayed significant faecal microbiota and plasma metabolome alterations at SHIP-2 when the respective diseases were absent. *Clostridium XIVa*, which was increased in these individuals, also associated with an increased expression of those microbial pathways necessary for biosynthesis of saturated and unsaturated fatty acids, some of which have been described to promote hepatic lipogenesis[46 47] or are correlated to increased hepatic lipogenesis.[48] Therefore, fatty acids synthesised by *Clostridium XIVa* may increase hepatic accumulation of lipids, thus contributing to the development of FLD toward SHIP-3. Whether *Clostridium XIVa* could serve as a treatment target for prebiotic or probiotic therapy to reduce the incidence of FLD needs to be determined in future interventional trials.

The advantage of this investigation is the usage of longitudinal microbiome data using paired samples at two time points to overcome the limitations of cross-sectional investigations. Moreover, the extensive phenotype database of the SHIP study allowed to identify critical host factors for faecal microbiota stability. A disadvantage of the study is the use of population-based data. Only the faecal microbiota relationship with conditions or states that are frequent enough to generate sufficient statistical power can be analysed in population-based cohorts, whereas estimation of the effect of less common or rare disorders on microbiota stability and composition requires disease-specific patient cohorts.[49]

Here we present results from a large-scale faecal microbiota follow-up study over 5 years. With ageing, the amount of facultative pathogenic bacteria increased. However, ageing itself was not a predictor of faecal microbiota changes, whereas metabolic disorders were. Why some components of metabolic syndrome such as diabetes mellitus and FLD have a greater association than others with faecal microbiota instability and whether this is due to associated changes in gut immunology or variations in intraluminal nutrient supply needs to be investigated in human interventional trials, in which the respective disease process can actually be reversed. The association of increased instability of microbiota composition over time with metabolic syndrome is clearly suggestive of a direct causality. The observation, however, that participants who newly developed diabetes mellitus or FLD at the later time point, but 5 years earlier had no detectable manifestation of the disease, while already displaying significant changes in microbiota composition, clearly indicates that the relationship is more complex. On the other hand, a greater diversity of the intestinal microbiome, female sex and preserved exocrine pancreatic function were found to promote a greater long-term stability of the faecal microbiome and may therefore convey greater resilience against common, dysbiosis-associated disorders.

Author affiliations
[1]Department of Medicine A, University Medicine Greifswald, Greifswald, Germany
[2]Chair of Experimental Bioinformatics, TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising-Weihenstephan, Germany
[3]Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany
[4]Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Germany
[5]DZHK (German Centre for Cardiovascular Research), partner site, Greifswald, Germany
[6]Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Germany
[7]Institute for Community Medicine, University of Greifswald, Greifswald, Germany
[8]Department of Internal Medicine B, University of Greifswald, Greifswald, Germany
[9]Department of Medicine II, University Hospital, LMU Munich, Munich, Germany

Twitter Malte Rühlemann @mruehlemann

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study was approved by the local institutional review board, University Medicine Greifswald, Germany, registration numbers BB 39/08 and BB 122/13.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. All microbiome and phenotype data were obtained from the SHIP data management unit and can be applied for online through a data access application form (https://www.fvcm.med.uni-greifswald.de/dd_service/data_use_intro.php).

ORCID iDs
Fabian Frost http://orcid.org/0000-0002-1502-8888
Malte Rühlemann http://orcid.org/0000-0002-0685-0052
Corinna Bang http://orcid.org/0000-0001-6814-6151
Andre Franke http://orcid.org/0000-0003-1530-5811
Julia Mayerle http://orcid.org/0000-0002-3666-6459
Markus M Lerch http://orcid.org/0000-0002-9643-8263

REFERENCES

1. Sender R, Fuchs S, Milo R. Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol* 2016;14:e1002533.
2. Zierer J, Jackson MA, Kastenmüller G, *et al*. The fecal metabolome as a functional readout of the gut microbiome. *Nat Genet* 2018;50:790–5.
3. Rastelli M, Knauf C, Cani PD. Gut microbes and health: a focus on the mechanisms linking microbes, obesity, and related disorders. *Obesity* 2018;26:792–800.
4. Tilg H, Mathurin P. Altered intestinal microbiota as a major driving force in alcoholic steatohepatitis. *Gut* 2016;65:728–9.
5. Chu H, Duan Y, Yang L, *et al*. Small metabolites, possible big changes: a microbiota-centered view of non-alcoholic fatty liver disease. *Gut* 2019;68:359–70.
6. Yatsunenko T, Rey FE, Manary MJ, *et al*. Human gut microbiome viewed across age and geography. *Nature* 2012;486:222–7.
7. Frost F, Storck LJ, Kacprowski T, *et al*. A structured weight loss program increases gut microbiota phylogenetic diversity and reduces levels of Collinsella in obese type 2 diabetics: a pilot study. *PLoS One* 2019;14:e0219489.
8. Vangay P, Johnson AJ, Ward TL, *et al*. US immigration Westernizes the human gut microbiome. *Cell* 2018;175:962–72. e10.
9. Odamaki T, Kato K, Sugahara H, *et al*. Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study. *BMC Microbiol* 2016;16:90.
10. Tamburini S, Shen N, Wu HC, *et al*. The microbiome in early life: implications for health outcomes. *Nat Med* 2016;22:713–22.
11. Völzke H, Alte D, Schmidt CO, *et al*. Cohort profile: the study of health in Pomerania. *Int J Epidemiol* 2011;40:294–307.
12. Mayerle J, den Hoed CM, Schurmann C, *et al*. Identification of genetic loci associated with Helicobacter pylori serologic status. *JAMA* 2013;309:1912–20.
13. Frost F, Kacprowski T, Rühlemann M, *et al*. Helicobacter pylori infection associates with fecal microbiota composition and diversity. *Sci Rep* 2019;9:20100.
14. Frost F, Kacprowski T, Rühlemann M, *et al*. Impaired exocrine pancreatic function associates with changes in intestinal microbiota composition and diversity. *Gastroenterology* 2019;156:1010–5.
15. Frost F, Kacprowski T, Rühlemann MC, *et al*. Functional abdominal pain and discomfort (IBS) is not associated with faecal microbiota composition in the general population. *Gut* 2019;68:1131.1–3.
16. Callahan BJ, McMurdie PJ, Rosen MJ, *et al*. DADA2: high-resolution sample inference from illumina amplicon data. *Nat Methods* 2016;13:581–3.
17. Douglas GM, Maffei VJ, Zaneveld JR, *et al*. PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol* 2020;38:685–8.
18. Levey AS, Stevens LA, Schmid CH, *et al*. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009;150:604–12.
19. Luedemann J, Schminke U, Berger K, *et al*. Association between behavior-dependent cardiovascular risk factors and asymptomatic carotid atherosclerosis in a general population. *Stroke* 2002;33:2929–35.
20. Winkler G, Döring A. Validation of a short qualitative food frequency list used in several German large scale surveys. *Z Ernährungswiss* 1998;37:234–41.
21. R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2017. https://www.R-project.org/
22. Oksanen J, Blanchet FG, Friendly M, *et al*. Vegan: community ecology package. R package version 2.4-2, 2017. Available: https://CRAN.R-project.org/package=vegan
23. Caspi R, Billington R, Fulcher CA, *et al*. The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res* 2018;46:D633–9.
24. Wang J, Kurilshikov A, Radjabzadeh D, *et al*. Meta-analysis of human genome-microbiome association studies: the MiBioGen Consortium initiative. *Microbiome* 2018;6:101.
25. Thingholm LB, Rühlemann MC, Koch M, *et al*. Obese individuals with and without type 2 diabetes show different gut microbial functional capacity and composition. *Cell Host Microbe* 2019;26:252–64. e10.
26. Tilg H, Cani PD, Mayer EA. Gut microbiome and liver diseases. *Gut* 2016;65:2035–44.
27. Cani PD, Van Hul M, diet M. Mediterranean diet, gut microbiota and health: when age and calories do not add up! *Gut* 2020;69:1167–8.
28. Wu H, Esteve E, Tremaroli V, *et al*. Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug. *Nat Med* 2017;23:850–8.
29. Olivares M, Neyrinck AM, Pötgens SA, *et al*. The DPP-4 inhibitor vildagliptin impacts the gut microbiota and prevents disruption of intestinal homeostasis induced by a Western diet in mice. *Diabetologia* 2018;61:1838–48.
30. Gomi H, Solomkin JS, Schlossberg D, *et al*. Tokyo guidelines 2018: antimicrobial therapy for acute cholangitis and cholecystitis. *J Hepatobiliary Pancreat Sci* 2018;25:3–16.
31. Lee C-C, Chang I-J, Lai Y-C, *et al*. Epidemiology and prognostic determinants of patients with bacteremic cholecystitis or cholangitis. *Am J Gastroenterol* 2007;102:563–9.
32. Li W, Chen H, Wu S, *et al*. A comparison of pyogenic liver abscess in patients with or without diabetes: a retrospective study of 246 cases. *BMC Gastroenterol* 2018;18:144.
33. Flores-Mireles AL, Walker JN, Caparon M, *et al*. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol* 2015;13:269–84.
34. Samonis G, Karageorgopoulos DE, Kofteridis DP, *et al*. Citrobacter infections in a general Hospital: characteristics and outcomes. *Eur J Clin Microbiol Infect Dis* 2009;28:61–8.
35. Pedersen C, Ijaz UZ, Gallagher E, *et al*. Fecal Enterobacteriales enrichment is associated with increased in vivo intestinal permeability in humans. *Physiol Rep* 2018;6:e13649.
36. Cani PD, Amar J, Iglesias MA, *et al*. Metabolic endotoxemia initiates obesity and insulin resistance. *Diabetes* 2007;56:1761–72.
37. Cani PD, Osto M, Geurts L, *et al*. Involvement of gut microbiota in the development of low-grade inflammation and type 2 diabetes associated with obesity. *Gut Microbes* 2012;3:279–88.
38. Cani PD, Bibiloni R, Knauf C, *et al*. Changes in gut microbiota control metabolic endotoxemia-induced inflammation in high-fat diet-induced obesity and diabetes in mice. *Diabetes* 2008;57:1470–81.
39. Cani PD, Possemiers S, Van de Wiele T, *et al*. Changes in gut microbiota control inflammation in obese mice through a mechanism involving GLP-2-driven improvement of gut permeability. *Gut* 2009;58:1091–103.
40. Riedel C-U, Foata F, Philippe D, *et al*. Anti-inflammatory effects of bifidobacteria by inhibition of LPS-induced NF-kappaB activation. *World J Gastroenterol* 2006;12:3729–35.
41. Cani PD, Neyrinck AM, Fava F, *et al*. Selective increases of bifidobacteria in gut microflora improve high-fat-diet-induced diabetes in mice through a mechanism associated with endotoxaemia. *Diabetologia* 2007;50:2374–83.
42. Imhann F, Bonder MJ, Vich Vila A, *et al*. Proton pump inhibitors affect the gut microbiome. *Gut* 2016;65:740–8.
43. Levine JM, D'Antonio CM. Elton revisited: a review of evidence linking diversity and invasibility. *Oikos* 1999;87:15.
44. Horvat S, Rupnik M. Interactions Between *Clostridioides difficile* and Fecal Microbiota in *in Vitro* Batch Model: Growth, Sporulation, and Microbiota Changes. *Front Microbiol* 2018;9:1633.
45. Ahuja M, Schwartz DM, Tandon M, *et al*. Orai1-mediated antimicrobial secretion from pancreatic acini shapes the gut microbiome and regulates gut innate immunity. *Cell Metab* 2017;25:635–46.
46. Ducheix S, Montagner A, Polizzi A, *et al*. Dietary oleic acid regulates hepatic lipogenesis through a liver X receptor-dependent signaling. *PLoS One* 2017;12:e0181393.
47. Yamada S, Kamada N, Amiya T, *et al*. Gut microbiota-mediated generation of saturated fatty acids elicits inflammation in the liver in murine high-fat diet-induced steatohepatitis. *BMC Gastroenterol* 2017;17:136.
48. Lee JJ, Lambert JE, Hovhannisyan Y, *et al*. Palmitoleic acid is elevated in fatty liver disease and reflects hepatic lipogenesis. *Am J Clin Nutr* 2015;101:34–43.
49. Cani PD. Human gut microbiome: hopes, threats and promises. *Gut* 2018;67:1716–25.