Supplementary Materials

Methods

16S microbiome data processing

The rationale behind the selection of 16S rRNA processing pipeline was described previously[12]. In short, the divergence in 16S rRNA gene domains between cohorts makes the OTU-level analysis impossible, while the use of direct taxonomic classification of the reads and of an up-to-date reference database allowed us to achieve better between-domain concordance of taxonomic composition and a higher mapping rate.

The participating cohorts varied in their sample collection protocol, selection of DNA purification kits used to extract DNA from fecal samples, the 16S domain selected for PCR (Table S1), read length, depth, post-sequencing quality control (QC) and the software used to merge tags of paired-end sequencing. After the processing of QC-filtered merged reads, all cohorts used the standardized 16S processing pipeline (https://github.com/alexa-kur/miQTL_cookbook) that uses SILVA release 128[57] as a reference database, with truncating the taxonomic resolution of the database to genus level.

First, all samples were rarefied to 10,000 reads using a predefined random seed to allow for rarefaction reproducibility. Samples with less than 10,000 reads were discarded. Second, RDP classifier v.1.12[58] was used to bin the reads to a reference database. For each taxonomic level, the posterior probability of 0.8 was used as a cutoff to bin each read to the corresponding taxon. The posterior cutoff probability was traced for each taxonomic level separately. For example, if the posterior probability passed the cutoff on the family level but not on the genus level, the read was binned to taxonomy on the family level (all corresponding upper taxonomic levels) and discarded on the genus level and assigned to special "NOTAX_genus" pseudo-taxon to maintain data compositionality.

To characterize the contribution of cohort-wise metadata (16S domain, DNA extraction method, cohort ethnicity, lysis temperature and type of lysis buffer) on the microbiome

composition, we used a distance-based redundancy analysis (dbRDA) test in which each cohort represented a sample, while variables represented mean abundances of genera in corresponding cohort. The association of metadata with richness was performed by merging individual richness data and performing multivariate linear regression analysis.

The alpha diversities, including Shannon, Simpson and inverse Simpson indices, were calculated on genus level, with non-adjusted, non-transformed taxa counts. For all other analyses, the taxonomic counts of non-zero samples were natural log-transformed and adjusted for potential covariate effects using linear regression. The list of covariates used in the regression models varied between cohorts, but always included sex, age, genetic principal components (PC) calculated on non-imputed genetic data (3 PCs for monoethnic cohorts, 10 PCs for multiethnic cohorts and 5 PCs for the HCHS/SOL cohort as a multi-ethnic population of different but closely related ethnicities, see Cohort descriptions) and cohort-specific potential microbiome batch effects, if applicable. Variables such as the length of time in in non-frozen storage, the 16S sequencing batch, etc. were also included. The residuals of the adjustment were then scaled and centered to the mean=0 and SD=1. For each cohort, only the taxa present in more than 10% of the samples were included in quantitative microbiome trait loci (mbQTL) mapping, whereas taxa present in more than 10% but less than 90% of the samples were included in binary trait loci (mbBTL) mapping (Table S3). The mbQTL and mbBTL approaches are described below.

In the analysis of microbiome composition heterogeneity, the cohorts SHIP/SHIP-TREND and GEM_HCE_v12/GEM_HCE_v24/GEM_HCE_ICHIP were merged to SHIP and GEM respectively, as they were analyzed with exactly the same protocols and in the same laboratories. In the microbiome-genetics analysis, these five cohorts were included individually as they differed in the genotype arrays and/or general populations represented.

Genetic data processing

Despite the difference in genomic array platforms, the majority of the cohorts used similar procedures for imputation, as well as post-imputation filtering steps. Twenty-four out of 25 cohorts used the Michigan Imputation Server (https://imputationserver.sph.umich.edu/index.html) for imputation, using the HRC 1.0 or 1.1 reference panel[59]. Due to the restrictions in manipulating data, the PNP study used an in-house pipeline for imputation instead; it used IMPUTE2[60,61] software (v.2.3.2) and 1000G reference

panel with addition population-matched genotypes of Jewish individuals[62]. The post-imputation cutoffs were the same for PNP and the other cohorts.

Post-imputation VCFs were transformed into TriTyper format and filtered using GenotypeHarmonizer software v.1.4.20 software[63]. The following cutoffs were applied for inclusion: minor allele frequency >0.05, pointwise imputation QC >0.4 and SNP-wise call rate filtering >0.95.

Heritability analysis

Heritability was calculated using data collected on 169 monozygotic (MZ) and 419 dizygotic (DZ) pairs of twins from the TwinsUK cohort (total of 1176 individuals). Twin-based heritability was calculated by fitting an ACE model on the transformed 16S rRNA data used for the GWAS using the OpenMx package (v.2.8.3) as previously described[6]. Prior to heritability estimation, the taxonomic abundance was normalized using inverse rank sum transformation. Since the NTR cohort comprised only monozygotic twins, the between-cohort heritability concordance was calculated as Pearson's correlation of intraclass correlation coefficient (ICC) for monozygotic twins. Pearson's correlation between NTR's and TwinsUK's ICCs was used to estimate the concordance. For mbQTLs, SNP-based heritability was calculated by LD score regression using 'LDSC' tool.

Microbiome GWAS analysis

The modified version of the eQTL mapping pipeline (https://github.com/molgenis/systemsgenetics/tree/master/eqtl-mapping-pipeline) was used to perform microbiome QTL (mbQTL) mapping[64].

The microbiome genome-wide association study (GWAS) was performed in three ways. First, we performed GWAS on three microbiome alpha diversity metrics (Shannon, Simpson and Inverse Simpson), using Spearman correlation between SNP dosages and alpha diversity metrics and adjusting for age, sex, technical covariates and genetic principal components.

Second, we used Spearman correlation to identify loci that affect the covariate-adjusted abundance of bacterial taxa, excluding samples with zero abundance (mbQTLs).

Third, we identified the loci associated with probability of presence vs absence of the bacterial taxon (mbBTLs). To reduce computational burden, the mbBTL analysis was performed in three stages. First, we calculated the Pearson correlation between SNP dosage and bacterial

presence encoded as 0/1, without adjusting for any covariate effect and using the previously mentioned eQTL mapping pipeline. Next, a weighted z-score meta-analysis was performed to calculate non-centrality for SNP-taxon association. Finally, all SNP-taxon pairs with meta P-value $<1x10^{-4}$ were recalculated using multiple logistic regression (R base package, versions from 3.2.0 to 3.5.1 depending on the group) with bacterial presence as an outcome and using SNP dosage along with the list of covariates as predictors. While the logistic regression provides more accurate estimates, it is also computationally intensive. We therefore opted for this three-stage procedure and limited the accurate estimation to the subset of SNPs showing some evidence of association in the first step analysis (Pearson correlation). All mbBTLs that reached nominal genome-wide significance threshold ($P<5x10^{-8}$) in logistic regression analysis had a Pearson correlation P-value more significant than $P<10^{-6}$, presuming the completeness of three-stage procedure in revealing genome-wide significant mbBTL using $P<10^{-4}$ cutoff at the first stage of analysis.

mbTL meta-analysis

Meta-analysis was performed using a weighted z-score method implemented in a modified version of BinaryMetaAnalyzer[64], a part of the eQTL mapping pipeline. Per-cohort Z-scores were calculated from Spearman correlation p-values using inverse normal transformation. For quantitative mbQTLs, the cohorts were weighted by the square root of the effective sample size (the number of samples having the bacterial taxon). For binary mbQTLs, the square root of the reported cohort size was used as a weighting for each study. The summary statistics generated for mbQTLs also include meta-effect sizes and standard errors. These were generated using the inverse variance weighted meta-analysis method performed on the per-cohort effect sizes and standard errors, backtracked from association Z-scores and minor allele frequencies using the strategy proposed and implemented by Zhu *et al*[65], where they also give the detailed derivation of the following equations:

$$\hat{b} = zS$$

$$S = \frac{1}{\sqrt{2p(1-p)(n+z^2)}}$$

Where, b is the estimated effect size, S is the estimated standard error, p is the allele frequency and n is the sample size.

Heterogeneity exploration analysis

Cross-study heterogeneity of the effects of genetic variants in the relative abundance of taxonomical units was assessed using Cochran's Q-test for heterogeneity[66], as implemented in METAL v2018-08-28[67], for all genome-wide significant variants ($P<5x10^{-8}$) found in our main analysis. To avoid reporting false-positive associations due to different study designs or data collection methods, we used a stringent threshold of $P<0.05$ to reject the null hypothesis of no heterogeneity. This threshold is conservative considering that several variants were tested simultaneously, and no correction for multiple testing was applied. In there was evidence of heterogeneity, a random effect model was also implemented to confirm the association results, using the metaphor R package v.2.0-0 (https://cran.r-project.org/web/packages/metafor/metafor.pdf).

Additionally, when there was evidence for heterogeneity of a SNP effect across cohorts, we implemented a meta-regression approach using the same package to assess whether variables such as age, ethnicity or sequenced region could explain the observed effect size heterogeneity.

SNP–age interaction analysis in LCT locus

To discover whether the association of functional SNPs in the *LCT* locus to the abundance of the *Bifidobacterium* genus varied between groups of adults and infants, we performed age–SNP interaction analysis in the GEM cohort, which comprises three sub-cohorts, each having a comparable number of individuals above and below puberty age. The age of 17 years old was selected to split the cohort into the age groups: adolescents or adults. Since the GEM cohort was composed of three sub-cohorts of different ethnic composition, we evaluated the interaction in both joint analysis and in each subcohort separately, using the following formula:

$$Bac = Sex + PC[1\text{-}3] + Age_{\text{group}} + Cohort + SNP_{\text{dos}} + SNP_{\text{HZ}} + SNP_{\text{GT}}:Age_{\text{group}}$$

where *Bac* is the log-transformed count of genus Bifidobacterium, *PC[1-3]* are three floats with the first 3 genetic PCs, *Cohort* is a batch variable that determines the cohort the sample belongs to, SNP_{dos} is a float-encoded dosage of alternative allele, SNP_{HZ} is a Boolean variable describing heterozygosity, SNP_{GT} is a genotype encoded as an unordered factor and Age_{group} is a two-level factor (above or below split level). The inclusion of a numeric dosage variable and a Boolean SNP_{HZ} variable allowed us to properly adjust for the recessive effect of the SNP on *Bifidobacterium* abundance without discarding SNP imputation uncertainty information embedded in SNP dosage.

The analysis was then repeated for each GEM subcohort separately, using the same model.

Association of mbTL-associated bacteria with host phenotypes

All the bacterial taxa found to be associated with mbTLs were correlated with 207 host phenotypes, including intrinsic host properties, diet, disease and medication information, in the LifeLines-DEEP (LLD) and FGFP cohorts. We used the Spearman correlation with Benjamini-Hochberg (BH)-adjustment for multiple testing to assess correlation between phenotypes and bacteria that had mbQTLs. For the taxa with mbQTLs, samples with zero abundance were truncated. For the taxa with mbBTLs, the abundance was transformed to a binary trait encoding presence/absence.

FUMA analyses of meta-analysis results

Functional mapping and annotation of 30 meta-analysis results were performed with FUMA (v1.3.5), an integrated web-based platform[68]. For each of 20 summary statistics from the mbQTL analyses, independent association signals were identified based on SNP P-values and LD between markers. Genome-wide significant loci and their boundaries were defined as non-overlapping genomic regions extending a LD window of $r^2 \geq 0.4$ (based on the 1000G EUR reference panel)[69] from the association signals with $P<5.0x10^{-8}$. Independent ($r^2<0.1$) lead SNPs from each locus were defined as those associated with the lowest P value from the region. Multiple risk loci were merged into a single genomic locus if the distance between their LD blocks was <250 kb.

Functional annotation of all candidate risk SNPs was obtained from different repositories integrated in FUMA. Furthermore, these functionally annotated SNPs were mapped to protein-coding genes using the following strategies. MbQTL analysis included: (1) positional mapping, with the maximum distance of 10 kb to protein-coding genes, and (2) eQTL mapping, using information from data repositories such as GTEx v7 and Blood eQTL browser (http://genenetwork.nl/bloodeqtlbrowser/)[70].

Since the procedure of mbBTL mapping provides accurate statistics for only the subset of SNPs (see Microbiome GWAS analyses paragraph), and we thus lack full summary statistics, only positional mapping was performed for mbBTLs, taking the protein-coding genes within 10 kb distance of the 10 leading SNPs per trait.

All mapped protein-coding genes were combined into one list for either mbQTL or mbBTL analysis, respectively, prior to performing GSEA integrated in FUMA. In further investigations, hypergeometric tests of enrichment of all mapped genes were performed not only in tissue-specific (differentially expressed) gene sets, but also in gene sets curated from various sources, e.g. MsigDB. We then reported all enriched gene sets (≥2) with FDR adjusted P-value <0.05.

PheWAS, genetic correlation and colocalization analysis

We performed the PheWAS look-ups in the summary statistics results of 4,155 traits collected by the GWASATLAS[29] (http://atlas.ctglab.nl/, accessed on: 25-09-2019) database for the top SNPs per mbQTL locus revealed either by mbQTL or mbBTL mapping. GWASATLAS includes 600 traits from the UK Biobank and is enriched with extensive phenotypes on proteomics (n=1124 proteins), hematology (n=36), metabolomics (n=1145 metabolic features) and immune markers (n=241) studied across variable sample sizes. It also contains 1,009 GWASs performed prior to the UK Biobank effort, all categorized under 27 phenotype domains. Next, we tested if any of these 27 domains were enriched by the phenotypes associated to one of the SNPs of interest (using a liberal P-value threshold for SNP–phenotype association of 0.05), as compared to the expected distributions under the null hypothesis. In order to obtain the distributions under the null hypothesis, we selected matching 1000 SNPs for each top SNP using SNPSNAP[71] matched by allele frequency, gene density, number of LD pairs and distance from the closest gene.

We then extracted corresponding results from the GWASATLAS for the matched 30,000 SNPs (1000 matching SNPs per each top mbTL SNP). The enrichment of each domain was tested by comparing the proportions of observed and expected significant results for the SNP of interest, by using the prop.test function in R. This resulted in one-sided P-values and odds ratios. Seven domains (Aging, Body structures, Connective tissue, Ear-Nose-Throat, Infection, Muscular and Social Interactions) that included fewer than 20 GWAS tables were excluded from the enrichment tests, resulting in 20 domains. We used a conservative Bonferroni-based P-value threshold of 8.33×10^{-5} for the enrichment testing, accounting for 20 domains and a total of 30 mbTL top SNPs coming from both the mbQTL and mbBTL mapping. In addition, we performed gene-based PheWAS look-ups in the GWASATLAS for the candidate genes of interest within 250 kb around the association peaks, as defined by the FUMA algorithms. This resulted in 146 genes to be tested.

The genetic correlation was estimated between Bifidobacterium and its PheWAS-related traits from Table S12 following a LD-score regression approach[72] using the 'ldsc' tool. For testing colocalization of the PheWAS signals, we used the approximate Bayes factor approach as implemented by the "coloc.abf" function from the "coloc" library in R[73], using genetic variants within ±250 kb around the top signals.

Mendelian randomization analysis

Mendelian randomization (MR) analyses were performed in R using TwoSampleMR package[33]. Causality direction was tested between the microbiome (mbQTLs from MiBioGen results) and two data types: (1) autoimmune, cardiovascular, metabolic (including weight-related phenotypes) and psychological diseases (GWAS summary statistics obtained from MRBase[33]), known to be associated with microbiome composition[3,4,41,74–79], and (2) 42 nutritional phenotypes and alcohol intake frequency from the UK Biobank round 2[80] (http://www.nealelab.is/uk-biobank/).

For MR analyses, the combined meta-effects and standard errors from inverse variance meta-analysis were used.

To test if a complex trait affected microbiome composition, we selected independent genetic variants associated with complex traits at the genome-wide significant level ($P<5x10^{-8}$) and used these as instruments in our MR analyses. For complex diseases, Odd Ratios (OR) and C.I., we transformed to the effect sizes and standard errors using built-in function of TwoSampleMR package. To test if microbiome changes were causally linked to complex traits, we first confined ourselves to bacteria with genome-wide significant QTLs ($P<5x10^{-8}$), and for these we selected all SNPs with a less stringent cut-off of $P<1x10^{-5}$ in our MR analyses as instruments. This strategy was used to increase the number of SNPs available in order to perform sensitivity analyses, as shown previously[47]. Independent SNPs were selected as instrumental variables based on $r^2 < 0.001$ in 1000G EUR data, and this was done within the TwoSampleMR package. When no shared SNPs between exposure and outcome were available, proxies from the 1000G EUR data ($r^2 > 0.8$) were added. We kept only the results based on at least three shared SNPs. The MR causality test was performed using the Wald ratio, and these Wald ratios were meta-analyzed using the inverse variance weighted (IVW) method[81]. We also estimated the causality using several additional methods: weighted mode method[82], which provides an

alternative approach to IVW; MR-Egger[83], which estimates the degree of horizontal pleiotropy in the data; and MR PRESSO[84], which estimates the pleiotropy and corrects for it by removing outliers from the inverse variance weighted model. We also assessed heterogeneity of the results obtained using Cochran's Q statistics[81] and by leave-one-out analyses[33]. We estimated instrument variable (IV) strengths using F statistics: the amount of variance explained by IVs was calculated for each exposure by using TwoSampleMR package (get_r_from_lor function) for binary traits and VPE as defined in Shi *et al*[85]. F statistics was then calculated as $\frac{r^2 * (N - 1 - k)}{(1 - r^2) * k}$, where r^2 is the variance explained, N is the sample size and k is the number of IVs. We kept the results the conventional threshold of F statistics >10[86].

After performing the MR tests, we excluded duplicated GWAS traits, as the same phenotype is often studied in multiple GWAS. To remove the duplicates, we kept the study with the largest sample size among all the tested GWAS studies for each trait.

After excluding duplicates and tests performed with weak instruments (F statistics <10), we then applied a BH correction for multiple testing to the results obtained from the IVW MR test, and subsequently used a stringent filtering procedure on the significant results to avoid false-positives. Specifically, we removed the MR results that were based on fewer than three SNPs and thus could not be further investigated with sensitivity analyses. We also removed MR results that were not supported by other MR tests (weighted mode method P >0.05, MR PRESSO P>0.05), or showed substantial pleiotropy or heterogeneity as estimated by MR-Egger (MR-Egger intercept P<0.05) or MR PRESSO outliers-adjusted test (P>0.05), as well as those where leave-one-out analysis identified one SNP driving the signal (all but one leave-one-out configurations had P<0.05). Of note, MR-Egger slope – representing the causal estimate – was not used as a filtering step given the reduced power to detect causal effects. It is also worth noting that for all but one of the reported MR results that passed all the filters above, the MR-Egger slope p-value was greater than 0.05, therefore an MR-Egger intercept P<0.05 cannot be used to exclude presence of pleiotropy. Even though many of our MR-Egger intercept results provided little evidence of directional pleiotropy, it is worth noting that a P<0.05 cannot exclude the presence of pleiotropy and requires further understanding of the biological mechanisms underpinning the relationship between genetic variation, the gut microbiome and health outcomes. To exclude more complex causality scenarios, we also removed those results for

which the reverse MR p-value was below 0.05. Of note, the identified causal relationship with the microbiome feature class Actinobacteria (as exposure) and ulcerative colitis (outcome) showed a consistent effect direction when using only the only one genome-wide significant SNP, but with wider confidence interval (OR=0.40 [95% CI: 0.22-0.71] $P_{nominal}$=0.002).

Cohort Descriptions

The total of 25 cohorts and 18,473 participants of different ethnicities and ages we participated in the microbiome GWAS analysis (Tables S1, S2).

BSPSPC (PopGen)

The PopGen cohort is a population-based cohort from the area around Kiel, Schleswig-Holstein, Germany. Participants were genotyped using the Affymetrix Genome-Wide Human SNP Array 6.0. Fecal samples of 714 individuals were collected by the participants themselves at home in standard fecal collection tubes and shipped to the study center where they were stored at -80°C until processing. DNA from fecal samples (app. 200 mg) was extracted using the QIAamp DNA stool mini kit automated on the QIAcube. Genotyping data generation, extraction of fecal DNA and sequencing of the V1-V2 variable region of the 16S rRNA gene and all data processing were performed at the Institute of Clinical Molecular Biology, Kiel, Germany.

CARDIA (Coronary Artery Risk Development in Young Adults Study)

Coronary Artery Risk Development in Young Adults Study (CARDIA) is a population-based prospective study of the evolution of cardiometabolic disease. African American and European American adults were recruited from four U.S. urban areas (Birmingham, AL; Chicago, IL; Minneapolis, MN; Oakland, CA in 1985-1986) (n=5,115, aged 18-30). They have subsequently been examined nine times. A microbiome study was initiated at the Year 30 follow-up examination (2015-2016) in a subset of participants (n=615) who had not taken antibiotics in the past month. Fecal DNA was extracted with the MoBio PowerSoil kit, and the V3-V4 region of the 16S rRNA gene was sequenced with Illumina MiSeq (2x300bp) at HudsonAlpha Institute for Biotechnology (Huntsville, AL, USA). A subset of cohort participants has been genotyped with the Affymetrix Genome-Wide Human SNP Array 6.0. After quality control and removing

participants with non-overlapping data on microbiome and host genetics, data from 114 African Americans and 257 European-Americans (total n=371) were available for analysis.

NeuroIMAGE+COMPULS

NeuroIMAGE+COMPULS is a cohort consisting of two studies, NeuroIMAGEII and COMPULS, and includes participants of Dutch ethnicity. The cohort represents a combination of adults/adolescents/children diagnosed with ADHD and healthy controls. The original design and updates of these studies have been described in detail[87,88]. The genotyping was performed by using the PsychChip array (BONN, Germany). Fecal DNA was extracted using a repeated bead beating and the Maxwell® 16 Instrument (Promega, Leiden, the Netherlands), as described previously[89]. Sequencing of bacterial 16S gene, domain V1-V2, was performed in GATC Biotech AG (Konstanz, Germany) using the Illumina HiSeq platform. The overlap between samples with genotyping and microbial 16S sequencing data yielded 133 samples (57 females, 76 males, 17(5) years old) for use in the microbiome GWAS analysis.

$COPSAC_{2010}$

The Copenhagen Prospective Studies on Asthma in Childhood 2010 ($COPSAC_{2010}$) cohort is a prospective mother-child cohort of 700 children and their families, recruited during week 24 of pregnancy[90]. The participants reside in and around Copenhagen, Denmark. The design builds upon the previous $COPSAC_{2000}$ cohort[91] and is based on detailed longitudinal clinical assessments of asthma, allergy, eczema, and other outcomes. Blood tests were taken from the infants at age of six months, and DNA was extracted from plasma. Genome-wide genotyping was performed using the Illumina OmniExpress-8 v1.4 and Exome BeadChip. Fecal samples were collected at visits to the clinic or at home by parents using detailed instructions, at ages 1 week, 1 month, and 1, 4, 5, and 6 years. For the present study, samples for age 4-6 years were used. Genomic DNA was extracted from the infants' samples using the PowerMag® Soil DNA Isolation Kit, and the V4 region of the 16S rRNA gene was amplified and sequenced on an Illumina MiSeq system, as previously described in detail[92]. At the relevant timepoint, we had both genotype and microbiome data for 380 children to include in this study, 73 of whom had taken antibiotics in the six months before the fecal sample date.

DanFunD (The Danish study of Functional Disorders)

The Danish Study of Functional Disorders (DanFunD) is a population-based cohort initiated to outline the epidemiology of functional somatic syndromes[93]. The study population comprises a random sample of 9,656 men and women aged 18-76 years from the general population, who were examined from 2011 to 2015. Genotyping using the Human OmniExpress Bead Array (Illumina Inc., San Diego, CA, USA) was conducted on human leukocyte DNA for the entire cohort. A subset of 2,464 participants volunteered to provide a fecal sample collected under standardized conditions. Microbial DNA extraction using the NucleoSpin Soil kit (Macherey-Nagel, Düren, Germany) and subsequent sequencing of the hypervariable region V4 of the bacterial 16S rRNA gene on an Illumina HiSeq 2500 platform was conducted at Beijing Genomics Institute (BGI Europe, Copenhagen, Denmark). In total, 2,396 samples passed the quality control for genotyping and 16S sequencing and were included in the GWAS.

FGFP (Flemish Gut Flora Project)

The Flemish Gut Flora Project is a population-based study cohort of 2,482 individuals from the Flanders region of Belgium. Blood and stool samples of volunteers were collected between June 2013 and April 2016. Genotyping was performed using the Human Core Exome arrays v1.0 and v1.1. Sampling kits were sent to the volunteers' homes and stored there at -18°C until collection and storage in the Raes Lab facilities at -80°C. DNA was extracted from the frozen fecal samples using the PowerMicrobiome RNA Isolation Kit as described in Falony et al[4]. Sequencing of the V4 region of the16S rRNA gene was carried out on the Illumina HiSeq platform. After quality control, 2,259 samples had genotype and 16S data (1,328 females, 896 males, mean age 52.3 yrs). FGFP procedures were approved by the medical ethics committee of the University of Brussels–Brussels University Hospital (approval 143201215505, 5/12/2012). A declaration concerning the FGFP's privacy policy was submitted to the Belgian Commission for the Protection of Privacy. Written informed consent was obtained from all participants.

FOCUS

The FoCus cohort is a population-based cohort from the area around Kiel, Schleswig-Holstein, Germany, and part of the competence network Food Chain Plus (FoCus, http://www.focus.uni-kiel.de/component/content/article/88.html). Participants were genotyped using the Infinium OmniExpressExome Array. All data generation and processing was performed at the Institute of Clinical Molecular Biology, Kiel, Germany, similar to the PopGen cohort.

GEM (The CCC GEM project)

The CCC GEM project is a prospective international research study that is designed to identify the potential triggers that contribute to the onset of Crohn's Disease. Since 2008, the GEM project has recruited over 5,000 healthy first-degree relatives of Crohn's Disease patients with an age range of 6-35 years. At the time of recruitment, participants were screened using a standardized questionnaire to exclude any history or symptoms of inflammatory bowel disease (IBD) or other gastrointestinal diseases. For the microbiome GWAS, we used data from participants recruited in Canada (n=1,115), United States (n=17) and Israel (n=111). Stool DNA was extracted using the QIAamp DNA Stool Mini Kit (Qiagen, Hilden, Germany). The V4 hypervariable region of bacterial 16S ribosomal RNA (16S rRNA) was sequenced using a MiSeq platform (Illumina Inc. San Diego, CA, USA) and primers 515F/806R[94]. The genotyping of the cohort was performed using the HumanCoreEXOME-12v1.1 chip (n=379), HumanCoreEXOME-24v1.0 chip (n=203) and both ImmunoChip and HumanCoreEXOME-12v1.1 chip (n=662) (Illumina, Inc. San Diego, CA, USA). Thus, in mbQTL mapping, the cohort was split into subcohorts GEM_v12, GEM_v24 and GEM_ICHIP respectively. Among subcohorts, GEM_v24 mostly comprises individuals of Israel ethnicity (70%, 61 Ashkenazi, 34 Sephardic, 18 other/unknown subethnicities), while the other two subcohorts are of a European ancestry. Only the sample from one member from each family enrolled in the project was included in the current microbiome GWAS study. After stringent quality control, as previously described[10], the overlap between samples with genotyping and microbial 16S sequencing data yielded 1,243 samples (676 females, 567 males, median age=19.0(8.03) yrs) for use in the microbiome GWAS analysis. None had used antibiotic in the three months before fecal collection.

The Generation R Study

The Generation R Study (GenR) is a population-based, prospective, multi-ethnic pregnancy cohort study from fetal life until young adulthood. It is conducted in the city of Rotterdam, the Netherlands[95].The genotyping of this cohort was performed using Illumina HumanHap 610K[96]. Stool sample collection started in 2012 and comprised 2,111 children. Fecal DNA was extracted using DiaSorin Arrow DNA (Isogen Life Science, De Meern, the Netherlands) with a bead-beating step. Sequencing of bacterial 16S gene, domain V3-V4, was performed in the Laboratory of Human Genetics at Erasmus MC Rotterdam, using the Illumina MiSeq platform[97]. After

stringent quality control, the overlap between samples with genotyping and microbial 16S sequencing data yielded 1,328 samples (656 females, 672 males, mean age 9.8(0.3) years) for use in the microbiome GWAS analysis. None had used antibiotics in the six months before fecal collection.

KSCS (Kangbuk Samsung Cohort Study)

The Kangbuk Samsung Cohort Study (KSCS) is a prospective cohort study to evaluate the natural history, prognosis, and genetic and environmental determinants of a wide range of health traits and diseases among Korean adults. There are two major cohort studies in Kangbuk Samsung Hospital: the KSCS and the Kangbuk Samsung Health Study (KSHS). The KSHS is a retrospective cohort study using de-identified data routinely collected during health screening visits from 2002 to present, includes standardized and high quality clinical, imaging and laboratory procedures and has information on multiple lifestyle and medical conditions. The KSCS is a prospective study that has now started to apply more strict and standardized procedures. They obtained informed consent for data linkage to national registries for death, cancer and medical utilization since 2011. Genotyping was conducted using the Illumina HumanCore BeadChips 12v in 2014 (n=2,040). Fecal samples were collected from 1,463 participants between June and September 2014. DNA extraction from fecal samples was performed within one month of storage using the MoBio PowerSoil ® DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA). Sequencing of bacterial 16S rRNA gene, domain V3-V4, was performed using the Illumina MiSeq platform (Illumina, San Diego, CA, USA). After quality control, 811 samples (319 females, 492 males, mean age 44.1 yrs) with overlapping genotype and 16S data were included in the microbiome GWAS.

LifeLines-DEEP (LLD)

The LifeLines-DEEP cohort (LLD) is a subcohort of the prospective LifeLines cohort from the northern provinces of the Netherlands (Groningen, Drenthe and Friesland) and includes participant of Dutch ethnicity. Blood and fecal samples of LifeLines-DEEP participants were collected between April and August 2013. The genotyping was performed using the Illumina ImmunoChip and Illumina Human CytoSNP-12 microarrays. Fecal DNA was extracted using the Qiagen AllPrep kit with bead-beating step. Sequencing of bacterial 16S gene, domain V4, was performed at the Broad Institute (Boston, USA) using the Illumina MiSeq platform. The overlap

between samples with genotyping and microbial 16S sequencing data yielded 875 samples (504 females, 371 males, mean age 45.4 (13.3) yrs) used for the microbiome GWAS analysis, of these 70 participants were PPI users and eight people used antibiotics in the six months previous to fecal collection.

METSIM

The METabolic Syndrome In Men (METSIM) cohort is a longitudinal population-based cross-sectional cohort comprising of 10,197 randomly selected non-diabetic Finnish men (aged from 45 to 73 years) who were examined in 2005-2010. Genotyping was performed using the Illumina Omni ExpressExome microarray. Microbial DNA was extracted from frozen fecal samples using the PowerSoil DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA) following the manufacturer's instructions. Amplification of the V4 hypervariable region of the 16S rRNA gene was done using the 515F and 806R primer and sequenced with the Illumina MiSeq platform at the University of California, Los Angeles. For the current microbiome GWAS study, we used a subset of the METSIM cohort consisting of 522 samples (mean age 61.91 (5.42) yrs) with overlapping genotyping and microbial 16S sequencing data.

MIBS

The Maastricht Irritable Bowel Syndrome (MIBS) cohort with biobank aims to identify subgroups of IBS according to phenotypical and genotypical characterization. At present, it includes 520 subjects with a clinical diagnosis of IBS according to the Rome III criteria (from primary-tertiary care) and 220 age- and gender-matched healthy controls. At baseline, all subjects completed an extensive questionnaire on demographics, lifestyle factors, medical history and medication use, as well as a 14-day symptom dairy, the GSRS, HADS, STAI, SF-36 and a food frequency questionnaire. In addition, blood (serum, (platelet poor) plasma, DNA), feces and exhaled air were collected. In subgroups, a rectal barostat and multisugar test for intestinal permeability was performed. All participants gave written informed consent. The project was approved by the Medical Ethics Committee of Maastricht University Medical Center+ and was registered in the US National Library of Medicine (http://www.clinicaltrials.gov, *NCT00775060*). For the present microbiome GWAS study, only controls (N=80) were included. Genotyping was performed using the Illumina ImmunoChip and Illumina Human CytoSNP-12 microarrays. Fecal DNA was extracted using the Qiagen AllPrep kit with bead-beating step.

Sequencing of bacterial 16S gene, domain V4, was performed at the Broad Institute (Boston, USA) using the Illumina MiSeq platform.

NGRC (NeuroGenetics Research Consortium)

The NeuroGenetics Research Consortium (NGRC) is a collaborative study of gene-environment-microbiome interaction on Parkinson's disease (PD). It is being conducted in the United States. GWAS was conducted in 2009 on DNA from whole blood on the Illumina HumanOmni1-Quad_v1-0_B array[98]. Stool and metadata were collected from a subset of participants in 2014. Fecal DNA was extracted using the MoBio PowerMag Soil DNA Isolation Kit (Optimized for KingFisher). 16S rRNA V4 amplicon was sequenced using the Illumina MiSeq platform[99]. For the microbiome GWAS study, only 133 control participants were used; they were free of neurodegenerative disease at a mean age of 71.9 (7.5) years old and 58% were female.

NTR (the Netherlands Twin Registry)

The Netherlands Twin Registry (NTR) collects data and biological samples on Dutch multiples and their family members[100]. The NTR samples included in the microbiome GWAS were collected for two separate studies: the first focused on the association between obesity and the gut microbiome and the other collected samples from family members and spouses. Genotyping was performed on the Affymetrix SNP 6.0, Affymetrix Axiom and Illumina GSA arrays. Fecal DNA was extracted using the Qiagen PowerSoil kit with the addition of the heating step of the Qiagen PowerFecal kit. The sequencing of the V4 domain of the 16S gene was performed using the Illumina MiSeq platform. DNA extractions and sequencing were performed at the Avera Institute for Human Genetics (Sioux Falls, SD, USA). One of each twin pair was randomly selected for inclusion in the GWAS analyses (156 twin pairs, 123 unrelated individuals, 279 individuals total). Both MZ twins were included for the ICC calculations between MZ twin pairs for comparison with heritability estimates (156 twin pairs). None of the participants reported using antibiotics within six months of fecal collection.

PNP (Personalized Nutrition Project)

The Personalized Nutrition Project is a large-scale nutrition initiative in Israel that aims to help people make food choices that would normalize their blood glucose level and therefore improve their health and well-being. The cohort has over 1,000 healthy individuals of Israeli ethnicity, living in Israel and aged between 18 and 70. The cohort consists of self-reported Ashkenazi

(n=508), North African (n=64), Middle Eastern (n=34), Sephardi (n=19), Yemenite (n=13) and 'admixed/other' (n=408) ancestries. The top two host genetic principal components are strongly associated with self-reported ancestry ($P<10^{-32}$ for both principal component 1 and principal component 2, Kruskal-Wallis test). Participants were genotyped using Illumina OMNI-EXPRESS arrays and they provided stool samples, which were collected using a swab or an OMNIGENE-GUT (OMR-200; DNA Genotek) stool collection kit. Metagenomic sequencing was performed on DNA extracted from the stool samples as well as 16S rRNA profiling by sequencing the V3-V4 region. 481 individuals were included in the current study.

PopCol (Population-based Colonoscopy)

Population-based Colonoscopy (PopCol) is a cohort study in Stockholm, Sweden, which includes a data-rich set of individuals with data available from bowel symptoms questionnaires, gastroenterology visits, and biospecimens (genotype and 16S sequencing from blood and stool samples, respectively)[101,102]. Genotyping was carried out using the Illumina HumanOmniExpressExome-8v1 arrays at the SciLifeLab NGI facility in Uppsala, Sweden. Fecal DNA was extracted from samples kept at -80°C using Qiagen QIAamp DNA Stool Mini Kits and analyzed using 16S rRNA gene amplicon sequencing (in the V1-V2 hypervariable region). This was performed on the Illumina MiSeq platform at the Institute of Clinical Molecular Biology (IKMB) in Kiel, Germany. After data merging and quality control, we used data from 134 individuals (83 females, 51 males, mean age 54.8(11.3) yrs) in the microbiome-GWAS. Of these, 6 PopCol participants were PPI users and 12 used antibiotics.

Rotterdam Study III

The Rotterdam Study (RS) is a prospective population-based cohort study established in 1990 to study determinants of disease and disability in Dutch adult/elderly individuals, aged ≥ 40 years. The original design and updates of this study have been described in detail[103]. RS consists of four sub-cohorts and comprises approximately 18,000 inhabitants of the Ommoord, a suburb of Rotterdam, the Netherlands. In the current microbiome GWAS, data from the Rotterdam Study III have been used. The genotyping was performed using the Illumina HumanHap 550K and 610K. The collection of fecal samples started in 2012 and includes 3,932 participants. Fecal DNA was extracted using DiaSorin Arrow DNA (Isogen Life Science, De Meern, the Netherlands) with a bead-beating step. Sequencing of bacterial 16S gene, domain V3-V4, was

performed in the Laboratory of Human Genetics at Erasmus MC Rotterdam, using the Illumina MiSeq platform[97]. After stringent quality control, the overlap between samples with genotyping and microbial 16S sequencing data yielded 1,220 samples (705 females, 515 males, mean age 57 (5.9) yrs) for use in the microbiome GWAS analysis. Of these, 260 participants used PPI and none used antibiotics in the six months before fecal collection.

SHIP (Study of Health in Pomerania)

The Study of Health in Pomerania (SHIP) is a prospective longitudinal population-based cohort study encompassing two independent cohorts: SHIP (N=4,308; baseline examinations 1997-2001) and SHIP-TREND (N=4,420; baseline examinations 2008 - 2012) [104]. Individuals were invited to the SHIP study center for computer-assisted personal interviews and extensive examinations. Follow-up investigations are scheduled at five-year intervals and have already been performed three times for SHIP and once for SHIP-TREND. For the microbiome GWAS project, data from the second SHIP wave (SHIP-2, 2008-2012), as well as the initial recruitment phase of SHIP-TREND, were used. Genotyping was performed using the Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA, USA) or the Infinium HumanOmni2.5 BeadChip (Illumina, San Diego, CA, USA) for SHIP and SHIP-TREND, respectively. Isolation of fecal DNA was done using the PSP Spin Stool DNA Kit (Stratec Biomedical AG, Birkenfeld, Germany). Fecal microbiota composition was determined based on the V1-V2 regions of the 16S rRNA gene on a MiSeq platform (Illumina) at the Institute of Clinical Molecular Biology (Christian Albrechts University of Kiel, Germany), as described before[105]. After comprehensive quality control, 1,901 datasets (1,043 females, 858 males, age 53.7 (14.0) yrs) with overlapping genotype and microbiome data were included in the current study. Of these, 149 individuals used PPI and 25 had antibiotics at the time of inclusion.

HCHS/SOL

The Hispanic Community Health Study/Study of Latinos (HCHS/SOL) is a prospective, population-based cohort study of 16,415 Hispanic/Latino adults (ages 18-74 years) who were selected using a two-stage probability sampling design from four US communities (Chicago, IL; Miami, FL; Bronx, NY; San Diego, CA)[106,107]. The genotyping of this cohort was performed with an Illumina custom array (15041502 B3), which consists of the Illumina Omni 2.5M array (HumanOmni2.5-8v1-1) plus ~150k custom SNPs, with the QC performed at HCHS/SOL

Genetic Analysis Center[108]. Stool sample were collected in the HCHS/SOL Gut Origins of Latino Diabetes (GOLD) ancillary study, which enrolled participants from the HCHS/SOL approximately concurrently with the second visit for HCHS/SOL. Fecal DNA was extracted with the Qiagen MagAttract PowerSoil DNA kit with both chemical and physical (i.e. bead-beating) means to release DNA, as described in Marotz et al[109]. Sequencing of bacterial 16S gene, domain V4, was performed in Rob Knight's lab at the University of California San Diego (San Diego, CA, USA) using the Illumina MiSeq platform. After stringent quality control, the overlap between genetically unrelated subjects with microbial 16S sequencing data yielded 1,097 samples (676 females, 421 males, age 57.2 (10.9) yrs) used in the microbiome GWAS analysis. Of these, 341 used medication including PPI for indigestion, heartburn, or stomach problems and 321 used antibiotics in the six months before the fecal collection.

TwinsUK

TwinsUK is a population-based cohort established in 1992 to study the genetic and environmental basis of a range of complex diseases and conditions in adult/elderly twins from the UK [110]. Genotyping was performed using HumanHap610Q on 5,654 volunteers, followed by imputation. Fecal samples were collected between 2010 and 2016 for 1,793 of the genotyped twins. DNA was extracted using PowerSoil - htp DNA isolation kit and the V4 region of the 16s rRNA gene was sequenced using the Illumina MiSeq platform at the Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, USA. One twin out of each pair was randomly excluded from the population of 1,793 individuals, leaving 1,205 volunteers (1,101 females and 104 males, age 61.5 (10.7) yrs) on which to conduct the microbiome GWAS analysis. Of these, 78 used PPIs and 62 had antibiotics 6 months prior to sampling.

Cohort Acknowledgements

CARDIA

The Coronary Artery Risk Development in Young Adults Study (CARDIA) is supported by contracts HHSN268201800003I, HHSN268201800004I, HHSN268201800005I, HHSN268201800006I, and HHSN268201800007I from the National Heart, Lung, and Blood Institute (NHLBI), United States. The microbiome study was funded by K01-NH127159, by the NIA Intramural Research Program, P30-DK056350, and UL1-TR002489. Genotyping was

funded as part of the NHLBI Candidate-gene Association Resource (N01-HC-65226) and the NHGRI Gene Environment Association Studies (GENEVA) (U01-HG004729, U01-HG04424, and U01-HG004446). This manuscript has been reviewed and approved by CARDIA for scientific content.

NeuroIMAGE+COMPULS

This study is funded by the EU's Horizon 2020 research and innovation program under Marie Skłodowska-Curie grant agreement No. 643051 (MiND) and grant agreement No. 728018 (Eat2BeNICE), and by the NIH project No. 1R01MH106655. This work is part of the research programme *Computing Time National Computing Facilities Processing Round pilots 2018* with project number *17666*, which is (partly) financed by the Dutch Research Council (NWO). This work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative. The generation and management of GWAS genotype data and microbiome data for the NeuroIMAGE+COMPULS was done at the Department of Human Genetics, Nijmegen, the Netherlands, under the supervision of Barbara Franke. The laboratory processing of stool microbiome samples for the NeuroIMAGE+COMPULS was executed by the Laboratory of Microbiology, Wageningen University, the Netherlands, under the supervision of Clara Belzer. We thank Remco Makkinje and Marlies Schimmel-Naber for their help in GWAS genotype data acquisition. We thank Marlies Schimmel-Naber and Saskia van der Marel for their help in administration and pre-laboratory processing of microbiome data, and Steven Aalvink and Sarita Dam for their help in DNA isolation and sequencing. All the authors are grateful to the participants and their families.

$COPSAC_{2010}$

We express our deepest gratitude to the children and families of the $COPSAC_{2010}$ cohort study for all their support and commitment. We acknowledge and appreciate the unique efforts of the COPSAC research team, and Martin S. Mortensen at the University of Copenhagen for extracting and sequencing of microbial DNA. COPSAC is funded by private and public research funds, including from the Lundbeck Foundation (grant no. R16-A1694); the Danish Ministry of Health (grant no. 903516); the Danish Council for Strategic Research (grant no. 0603-00280B); and the Capital Region Research Foundation. No pharmaceutical company was involved in the study.

DanFunD

The DanFunD staff are thanked for their commitment and perseverance in phenotyping study participants. We are grateful to Annemette Forman and Tine Lorentzen, lab technicians. The Lundbeck Foundation and the Tryg Foundation are thanked for their for financial support to the DanFunD microbiome study. The Novo Nordisk Foundation Center for Basic Metabolic Research is an independent Research Center at the University of Copenhagen, partially funded by an unrestricted donation from the Novo Nordisk Foundation (www.metabol.ku.dk). All other authors declare that there is no duality of interest associated with their contribution to this manuscript.

FGFP

We acknowledge the contribution of Flemish GPs and pharmacists to the data and sample collection. We thank all FGFP volunteers for participating in the project. The FGFP was funded with support from the Flemish government (IWT130359), the Research Fund–Flanders (FWO) Odysseus program (G.0924.09), the King Baudouin Foundation (2012-J80000-004), FP7 METACARDIS HEALTH-F4-2012-305312, VIB, the Rega Institute for Medical Research (Leuven), and KU Leuven. This study is partially funded by the FWO EOS program (30770923). RB is funded by FWO through a Postdoctoral Fellowship (1221620N). DH is supported by a Wellcome Trust Investigator Award (202802/Z/16/Z). KHW is supported by the Elizabeth Blackwell Institute for Health Research, University of Bristol, UK, and the Wellcome Trust Institutional Strategic Support Fund (204813/Z/16/Z). NJT is a Wellcome Trust Investigator (202802/Z/16/Z) supported by the University of Bristol NIHR Biomedical Research Centre (BRC-1215-20011) and works in the CRUK Integrative Cancer Epidemiology Program (C18281/A19169). NJT works in a unit that also receives funding from the MRC and the University of Bristol (MC_UU_12013/3, MC_UU_00011/6).

FOCUS/BSPSPC

We thank Ilona Urbach, Ines Wulf and Tonio Hauptmann of the IKMB microbiome laboratory and the staff of the IKMB sequencing facilities for excellent technical support. This work was supported by the Deutsche Forschungsgemeinschaft (DFG) Cluster of Excellence 'Inflammation at Interfaces' (http://www.inflammation-at-interfaces.de, nos. EXC306 and EXC306/2), the

Collaborative Research Center 1182, Origin and Function of Metaorganisms (www.metaorganism-research.com, no: SFB1182).

GEM

All authors disclose no potential conflicts (financial, professional, or personal) that are relevant to the manuscript. This study was supported by grants from Crohn's and Colitis Canada Grant #CCC-GEMIII, Canadian Institutes of Health Research (CIHR) Grant #CMF108031 and the Helmsley Charitable Trust. Williams Turpin is a recipient of a Postdoctoral Fellowship Research Award from the CIHR Fellowship/ Canadian Association of Gastroenterology (CAG)/ Ferring Pharmaceuticals Inc. and a fellowship from the Department of Medicine, Mount Sinai Hospital, Toronto. Juan Antonio Raygoza Garay holds a fellowship from the Department of Medicine, Mount Sinai Hospital, Toronto. Kenneth Croitoru is partially supported by a Canada Research Chair in Inflammatory Bowel Diseases. We thank the members of the CCC GEM Global Project Office: Ashleigh Goethel, Heather MacAulay, Stephanie Gosselin, Yurie Yamagishi, Clarissa Ratjen and all past team members for administrative support. The CCC GEM Project Research Consortium is composed of: Maria Abreu, Paul Beck, Charles Bernstein, Kenneth Croitoru, Leo Dieleman, Brian Feagan, Anne Griffiths, David Guttman, Kevan Jacobson, Denis O. Krause*, Karen Madsen, John Marshall, Paul Moayyedi, Mark Ropeleski, Ernest Seidman*, Mark Silverberg, Scott Snapper, Andy Stadnyk, Hillary Steinhart, Michael Surette, Dan Turner, Thomas Walters, Bruce Vallance, Alain Bitton, Maria Cino, and Dennis Cvitkovitch.(*deceased)

The recruitment site directors include Maria Abreu, Guy Aumais, Robert Baldassano, Charles Bernstein, Maria Cino, Lee Denson, Colette Deslandres, Wael El-Matary, Anne M. Griffiths, Charlotte Hedin, Hans Herfarth, Peter Higgins, Seamus Hussey, Hien Huynh, Jeff Hyams, Kevin Jacobson, David Keljo, David Kevans, Charlie Lees, David Mack, John Marshall, Jerry McGrath, Sanjay Murthy, Anthony Otley, Remo Panaccione, Nimisha Parekh, Sophie Plamondon, Graham Radford-Smith, Mark Ropeleski, Joel Rosh, David Rubin, Michael Schultz, Ernest Seidman, Corey Siegel, Scott Snapper, Hillary Steinhart, Alain Bitton, Martha Dirks, Brian Feagan, Herbert Brill, Sonia Michail, Hillel Noan, Jennifer Jones, Pierre Pare, Neil Leleiko, Fred Saibil, and Dan Turner.

Generation R

We gratefully acknowledge the contribution of children and parents, general practitioners, hospitals, midwives and pharmacies in Rotterdam. The generation and management of GWAS genotype data for the Generation R Study was done at the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Netherlands. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for their help in creating, managing and QC of the GWAS database. The musculoskeletal research of the Generation R Study is partly supported by the European Commission grant HEALTH-F2-2008-201865-GEFOS. The general design of Generation R Study is made possible by financial support from the Erasmus Medical Center, Rotterdam, the Erasmus University Rotterdam, the Netherlands Organization for Health Research and Development (ZonMw), the Netherlands Organization for Scientific Research (NWO), the Netherlands' Ministries of Health, Welfare and Sport and of Youth and Families. Additionally, the Netherlands Organization for Health Research and Development supported authors of this manuscript (ZonMw 907.00303, ZonMw 916.10159, ZonMw VIDI 016.136.361 to VWJ, and ZonMw VIDI 016.136.367 supporting CMG). VWJ received a Consolidator Grant from the European Research Council (ERC-2014-CoG-648916). JFF received funding from the EU's Horizon 2020 research and innovation program under grant agreement No 633595 (DynaHEALTH). The study was also supported by funding from the EU's Horizon 2020 research and innovation program (733206, LIFECYCLE). The generation and management of stool microbiome data for the Generation R Study was executed by the Human Genotyping Facility of the Genetic Laboratory in the Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands. We thank Nahid El Faquir and Jolande Verkroost-Van Heemst for their help in sample collection and registration, and Kamal Arabe, Hedayat Razawy and Karan Singh Asra for their help in DNA isolation and sequencing. Furthermore, we thank Jeroen Raes and Jun Wang (KU Leuven, Belgium) for their guidance in 16S rRNA profiling and dataset generation. Djawad Radjabzadeh was funded by an Erasmus MC mRACE grant "Profiling of the human gut microbiome".

KSCS

This study was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1B07050067) and by the Medical Research Funds from Kangbuk Samsung Hospital. We thank the participants and staff of the Kangbuk Samsung Cohort Study for their

collaboration, especially Yoosoo Chang and Seungho Ryu. The study was approved by the EUMC review board 2014-06-024 and KBSMC review board 2013-01-245. The computing resources were supported by the Global Science Experimental Data Hub Center (GSDC) Project and Korea Research Environment Open NETwork (KREONET) at the Korea Institute of Science and Technology Information (KISTI).

LLD

LifeLines DEEP: we thank participants and staff of the LifeLines DEEP cohort for their collaboration. We thank J. Dekens, M. Platteel, A. Maatman and J. Arends for management and technical support. This project was funded by the Netherlands Heart Foundation (IN-CONTROL CVON, grant no. 2012-03 to A.Z. and J.F.); by NWO (nos. NWO-VIDI 864.13.013 to J.F., NWO-VIDI 016.178.056 to A.Z., and NWO Gravitation Program ExposomeNL to A.K. and A.Z.), and by the European Research Council Starting Grant no. 715772 to A.Z., who also holds Rosalind Franklin Fellowship from the University of Groningen. A. D. was supported by a VENI grant from NWO, H2020-SC1-HBC-28-2019-LONGITOOLS and WCRF-2017/1641. D.V.Z. was supported by a VENI grant from NWO no. 194.006.

MIBS

We thank all participants and M. Hesselink for her support in data collection. Part of the data collection and analyses was funded by the Top Institute for Food and Nutrition (TIFN), Wageningen, the Netherlands.

METSIM

We thank all METSIM study participants. This study was funded by the Estonian Research Council grant (PUT1371), the EMBO Installation grant (No 3573), the US National Institutes of Health (NIH) grants (HL28481, HL30568 and HL123295), the Academy of Finland (321428), Sigrid Juselius Foundation, Finnish Foundation for Cardiovascular Research, and the Centre of Excellence of Cardiovascular and Metabolic Diseases, supported by the Academy of Finland.

NGRC

We thank the NeuroGenetics Research Consortium (NGRC) participants and our collaborators S. Factor, C. Zabetian and E. Molho. Funding from the National Institutes of Health (USA), Michael J Fox Foundation for Parkinson's Disease Research, and NIH Training Grant T32NS095775.

PNP

We thank the Segal and Elinav group members for discussions. ES is supported by the Crown Human, Genome Center and the Else Kroener Fresenius Foundation; DL Schwarz; JN Halpern; L Steinberg; and grants from the European Research Council and the Israel Science Foundation. DR received a Levi Eshkol PhD Scholarship for Personalized Medicine from the Israeli Ministry of Science.

POPCOL

PopCol data generation and analysis has been partially supported by funding from the Swedish Research Council (VR project 2017-02403) available to Mauro D'Amato.

Rotterdam Study III

The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organization of Scientific Research (NWO) Investments (no. 175.010.2005.011, 911-03-012). This study was funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO) project no. 050-060-810, and Memorabel (ZonMW project no. 733050814). We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera, Marjolein Peters and Carolina Medina-Gomez for help in creating the GWAS database, and Linda Broer and Carolina Medina-Gomez for the creation and analysis of imputed data. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam; the Netherlands Organization for Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly (RIDE); the Netherlands' Ministries of Education, Culture and Science, and of Health, Welfare and Sports; the European Commission (DG XII), and the Municipality of Rotterdam. The authors thank the study participants, the Rotterdam Study staff, and the general practitioners and pharmacists who participated. The generation and management of stool microbiome data for the Rotterdam Study (RSIII-2) was executed by the Human Genotyping Facility of the Genetic Laboratory, Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands. We thank Nahid El Faquir and Jolande Verkroost-Van Heemst for their help in sample collection and registration, and Kamal Arabe, Hedayat Razawy and Karan Singh Asra for help in DNA isolation and sequencing. Furthermore, we thank Jeroen Raes and Jun Wang (KU Leuven, Belgium) for guidance in 16S rRNA profiling and dataset generation.

Djawad Radjabzadeh was funded by an Erasmus MC mRACE grant “Profiling of the human gut microbiome”.

SHIP

SHIP is part of the Community Medicine Research network of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grant nos. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs, the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network ‘Greifswald Approach to Individualized Medicine (GANI_MED)’, funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide SNV typing data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthineers, Erlangen, Germany and the Federal State of Mecklenburg-West Pomerania. Further support was received from the RESPONSE-project (BMBF grant number 03ZZ0921E), PePPP-project (ESF/14-BM-A55_0045/16), and EnErGie-project (ESF/14-BM-A55-0010/18).

HCHS/SOL

The Hispanic Community Health Study/Study of Latinos is a collaborative study supported by contracts from the US National Heart, Lung, and Blood Institute (NHLBI) to the University of North Carolina (HHSN268201300001I / N01-HC-65233), University of Miami (HHSN268201300004I / N01-HC-65234), Albert Einstein College of Medicine (New York), (HHSN268201300002I / N01-HC-65235), University of Illinois at Chicago – HHSN268201300003I / N01-HC-65236 Northwestern Univ.), and San Diego State University (HHSN268201300005I / N01-HC-65237). The following US organizations have contributed to the HCHS/SOL by transferring funds to the NHLBI: National Institute on Minority Health and Health Disparities, National Institute on Deafness and Other Communication Disorders, National Institute of Dental and Craniofacial Research, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Neurological Disorders and Stroke, NIH Institution-Office of Dietary Supplements. The Genetic Analysis Center at the University of Washington was supported by NHLBI and NIDCR contracts (HHSN268201300005C AM03 and MOD03). Our study was also supported by grants from NIMHD (1R01MD011389-01) and NHLBI (1R01HL140976-01).

TwinsUK

TwinsUK is funded by the Welcome Trust, Medical Research Council, European Union, the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. SNP genotyping was performed by The Wellcome Trust Sanger Institute and National Eye Institute via NIH/CIDR. We thank Dr Julia K. Goodrich and Dr Ruth E. Ley and the Cornell technical team for generating the microbial data.

Supplementary figures

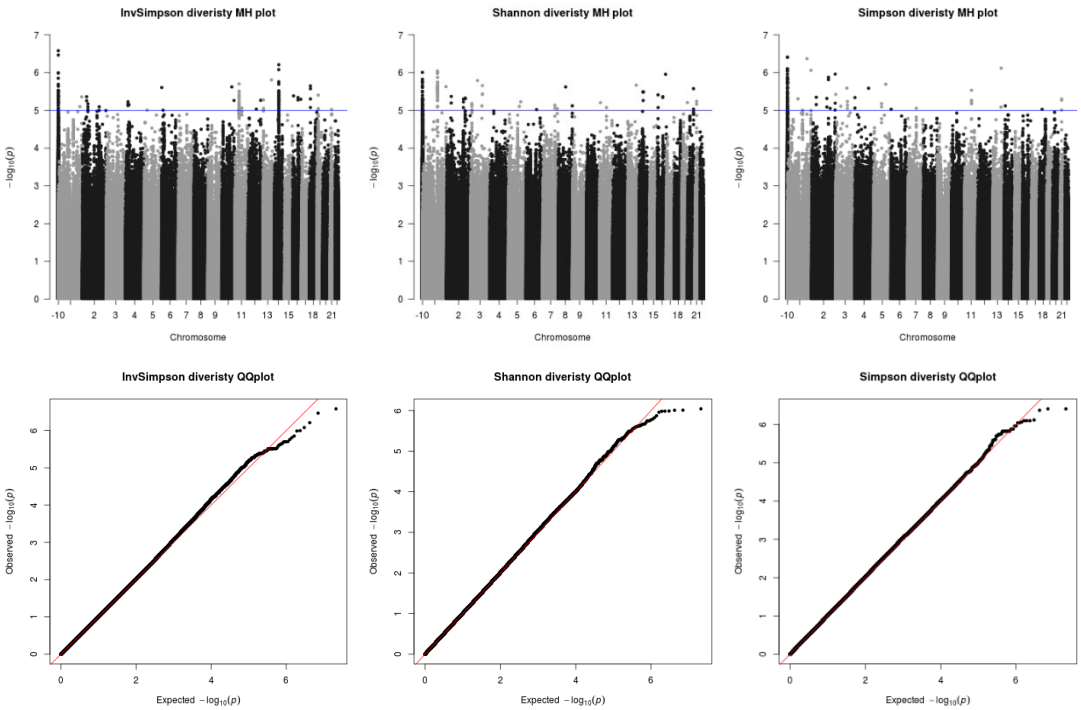
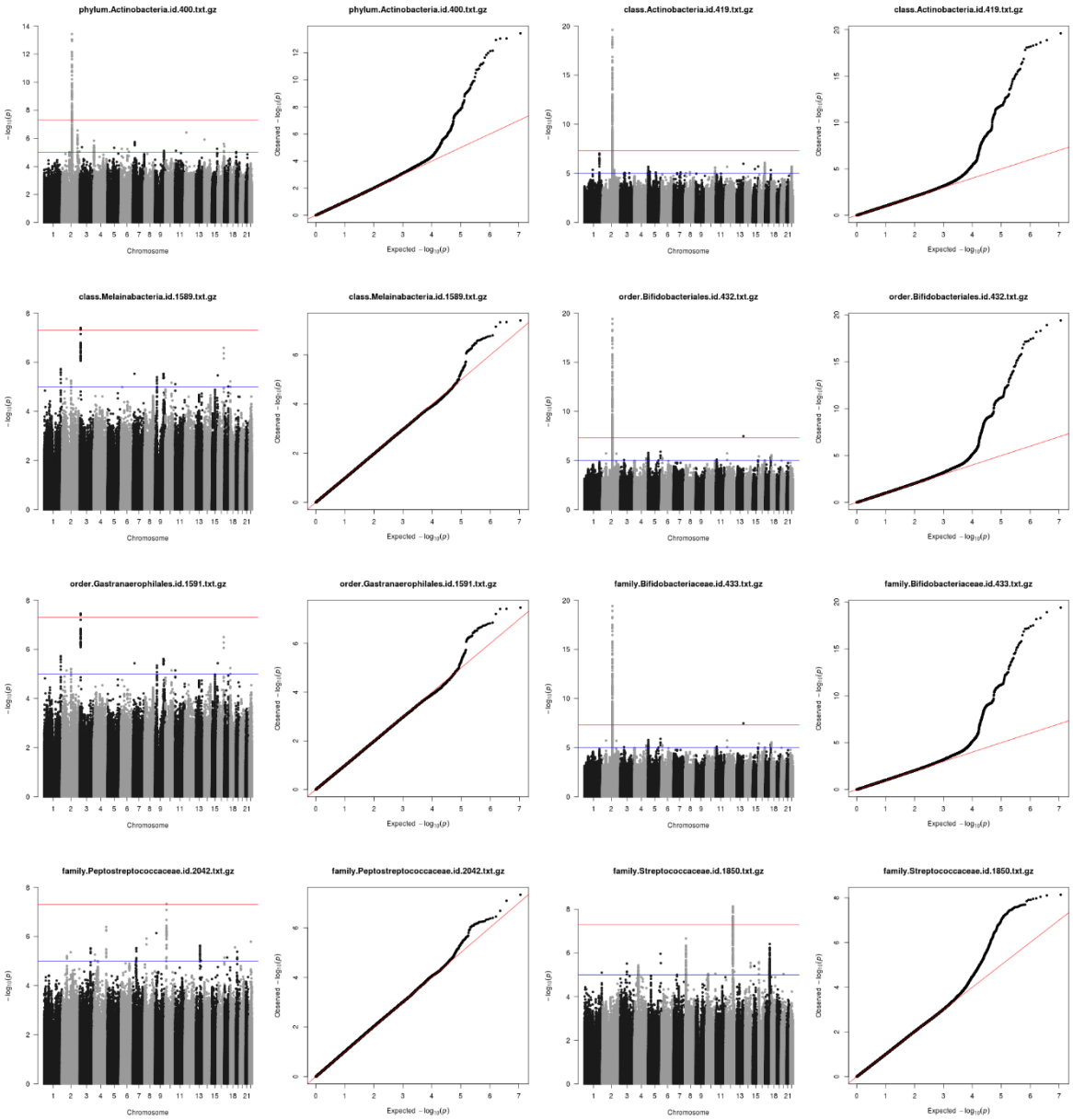
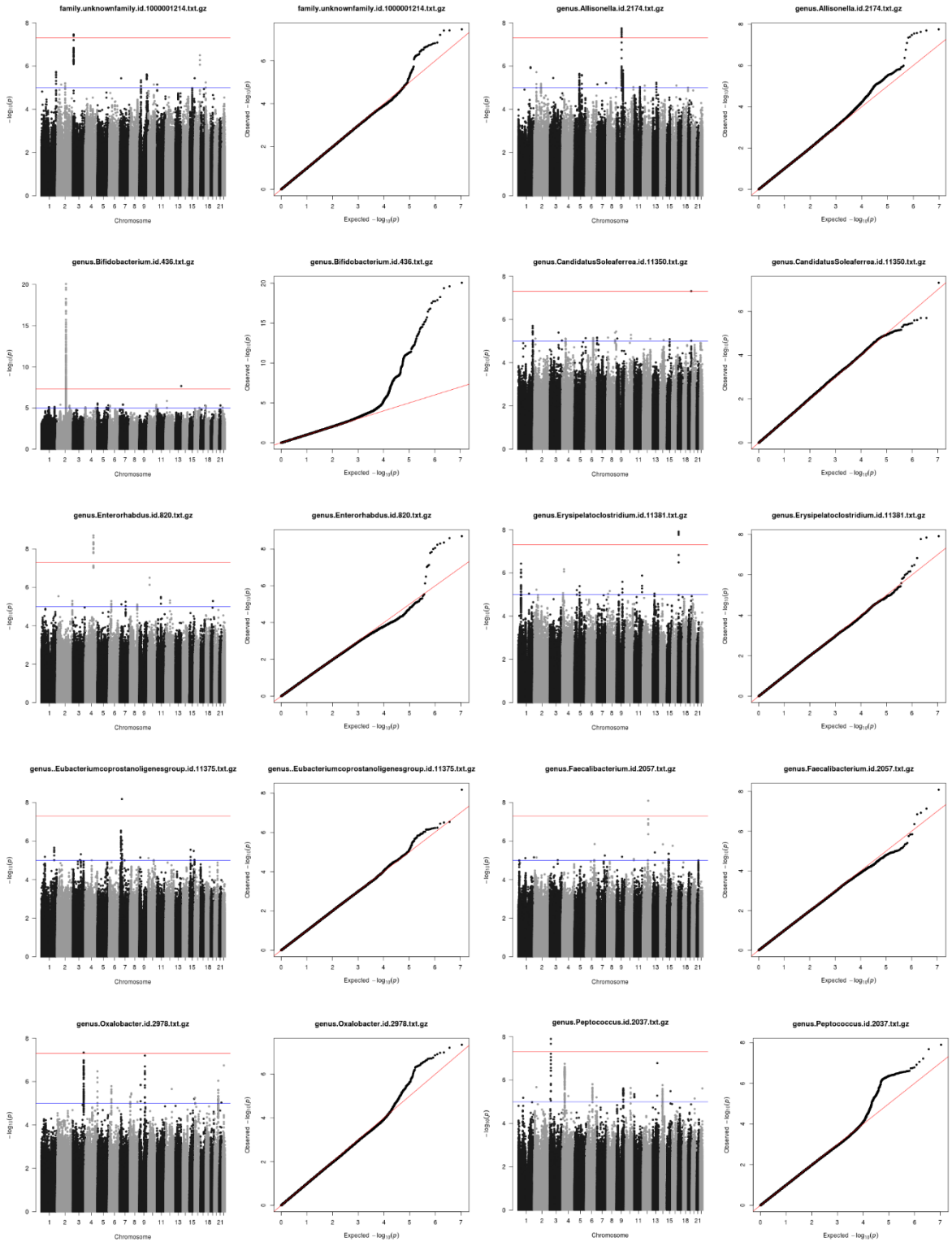


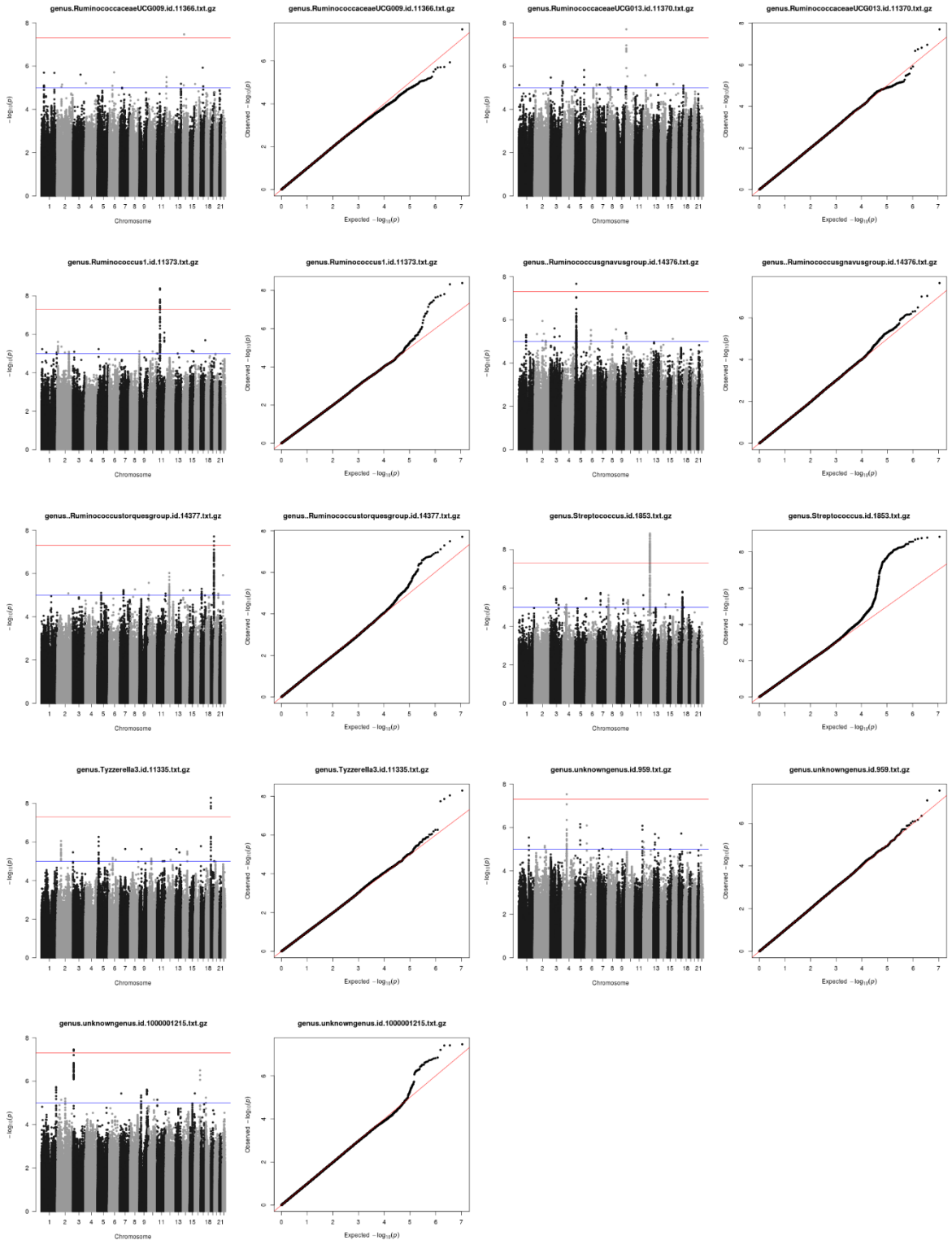
Figure S1. Manhattan plots (top) and QQ plots (bottom) of GWAS on alpha diversity metrics (Inverse-Simpson, Shannon, Simpson indices). The name of the trait is given in the title of each plot.

Figure S2. Manhattan plots and QQ plots for mbQTLs (placed in an order of taxonomic level, from high to low).




family.unknownfamily.id.1000001214.txt.gz
family.unknownfamily.id.1000001214.txt.gz
genus.Allisonella.id.2174.txt.gz
genus.Allisonella.id.2174.txt.gz
genus.Bifidobacterium.id.436.txt.gz
genus.Bifidobacterium.id.436.txt.gz
genus.CandidatusSoleaferrea.id.11350.txt.gz
genus.CandidatusSoleaferrea.id.11350.txt.gz
genus.Enterorhabdus.id.820.txt.gz
genus.Enterorhabdus.id.820.txt.gz
genus.Erysipelatoclostridium.id.11381.txt.gz
genus.Erysipelatoclostridium.id.11381.txt.gz
genus..Eubacteriumcoprostanoligenesgroup.id.11375.txt.gz
genus..Eubacteriumcoprostanoligenesgroup.id.11375.txt.gz
genus.Faecalibacterium.id.2057.txt.gz
genus.Faecalibacterium.id.2057.txt.gz
genus.Oxalobacter.id.2978.txt.gz
genus.Oxalobacter.id.2978.txt.gz
genus.Peptococcus.id.2037.txt.gz
genus.Peptococcus.id.2037.txt.gz
Chromosome
Expected −log10(p)
Observed −log10(p)

5


genus.RuminococcaceaeUCG009.id.11366.txt.gz
genus.RuminococcaceaeUCG009.id.11366.txt.gz
genus.RuminococcaceaeUCG013.id.11370.txt.gz
genus.RuminococcaceaeUCG013.id.11370.txt.gz
genus.Ruminococcus1.id.11373.txt.gz
genus.Ruminococcus1.id.11373.txt.gz
genus..Ruminococcusgnavusgroup.id.14376.txt.gz
genus..Ruminococcusgnavusgroup.id.14376.txt.gz
genus..Ruminococcustorquesgroup.id.14377.txt.gz
genus..Ruminococcustorquesgroup.id.14377.txt.gz
genus.Streptococcus.id.1853.txt.gz
genus.Streptococcus.id.1853.txt.gz
genus.Tyzzerella3.id.11335.txt.gz
genus.Tyzzerella3.id.11335.txt.gz
genus.unknowngenus.id.959.txt.gz
genus.unknowngenus.id.959.txt.gz
genus.unknowngenus.id.1000001215.txt.gz
genus.unknowngenus.id.1000001215.txt.gz
Chromosome
Expected −log10(p)
Observed −log10(p)

5

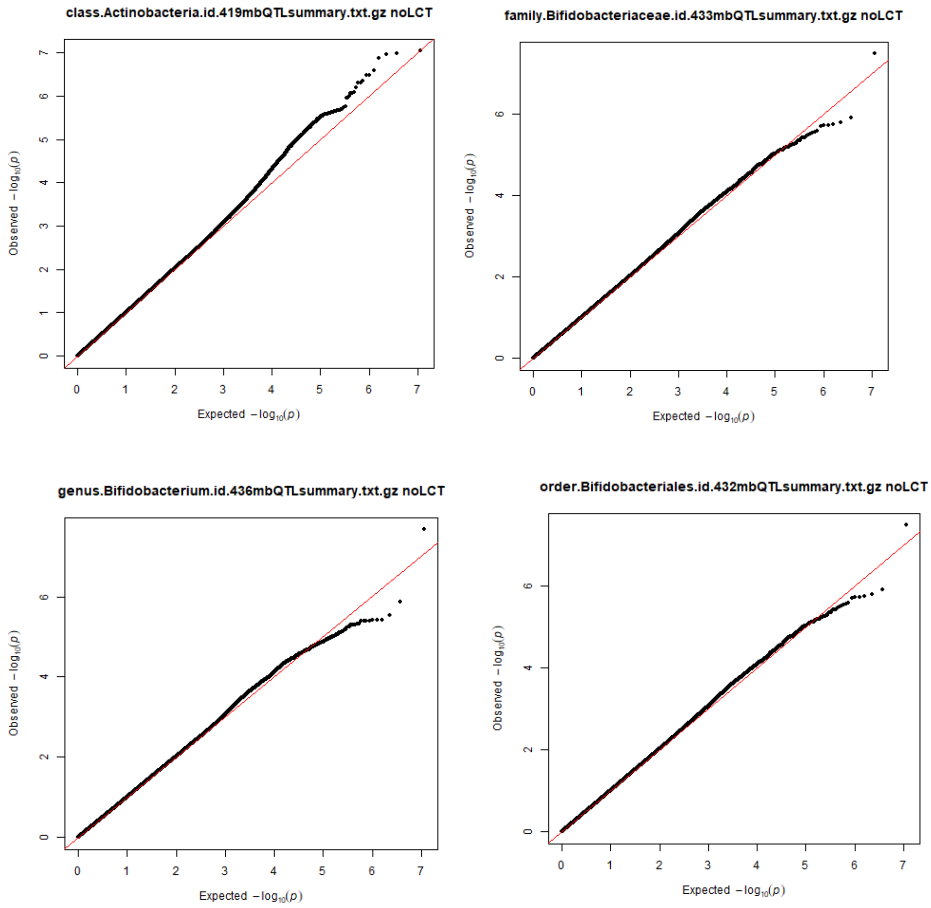


Figure S3. QQplots of genus Bifidobacterium and its related taxa of upper level, excluding *LCT* locus (2MB upstream and downstream of the top SNP, rs182549).

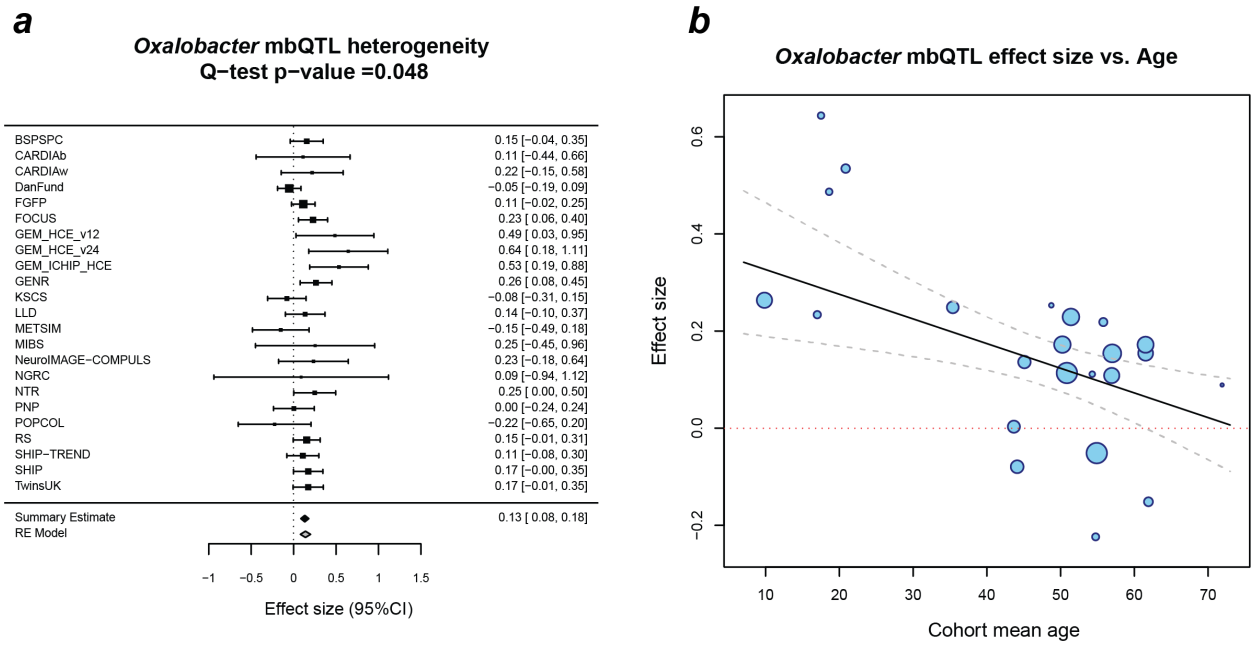


Figure S4. Heterogeneity analysis of *Oxalobacter* mbQTL. The estimations are given for the SNP rs4428215. **(A)** Forest plot of effect sizes across cohorts. Effect sizes were estimated from cohort Z-scores and allele frequencies, see Methods for more details. **(B)** Meta-regression of effect size of *Oxalobacter* mbQTL vs mean cohort age.

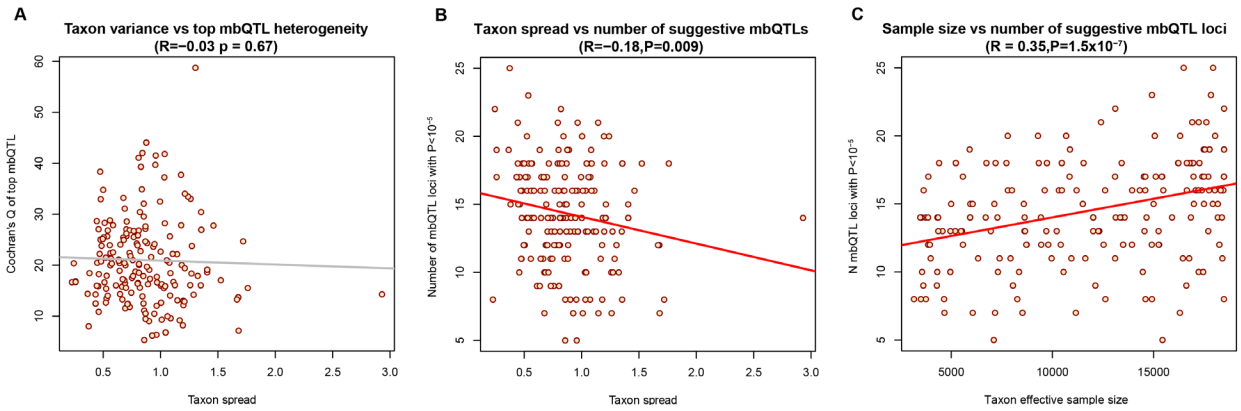
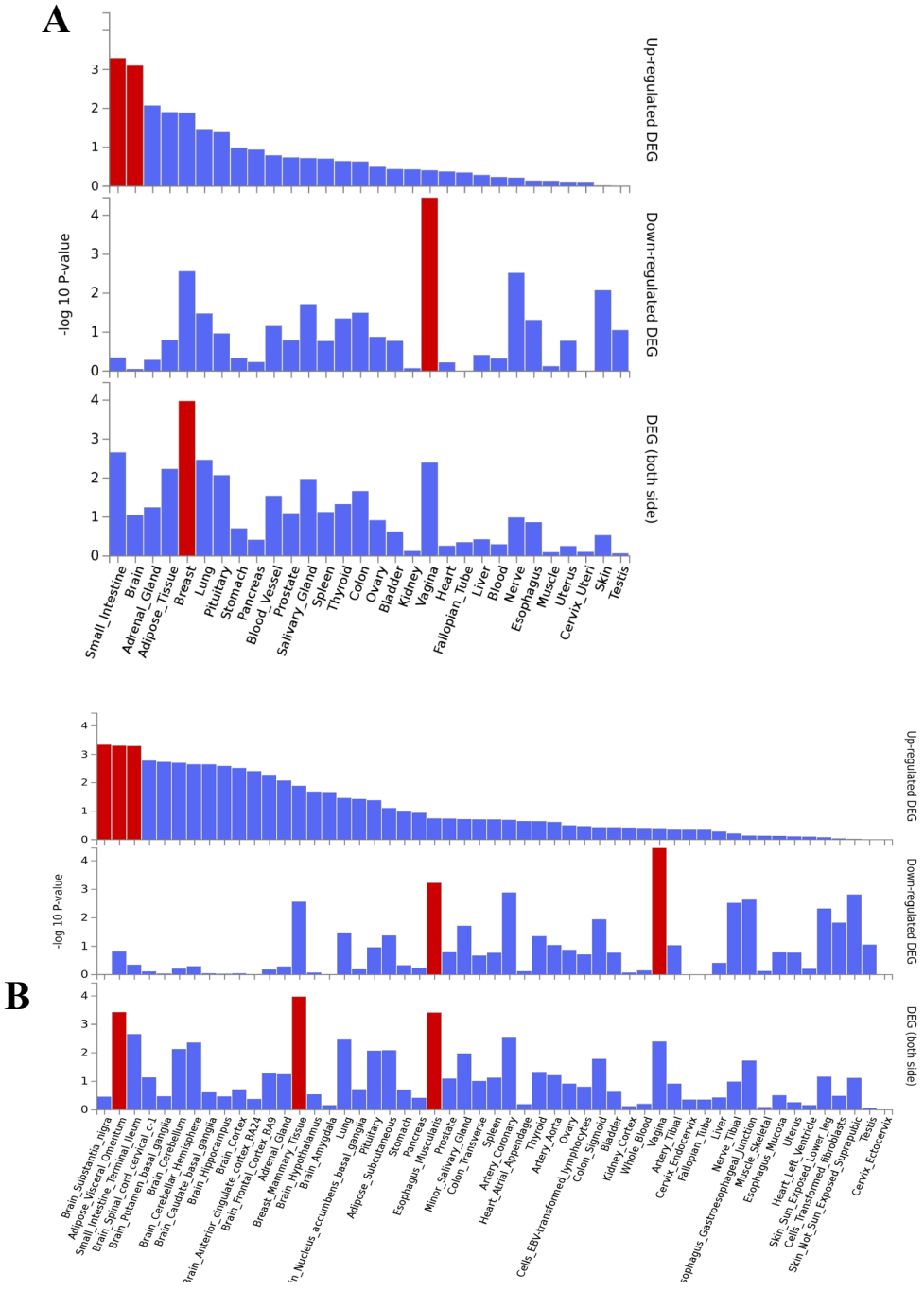


Figure S5. Effect of taxon sample size and taxon intercohort variance on the detectability and heterogeneity of mbQTLs. **(A)** The correlation of taxon intercohort spread (calculated as $SD_{cohortMean}/E_{cohortMean}$) and Cochran's Q of top mbQTL per taxon. The genus *Bifidobacterium* and its related taxa of upper levels (family *Bifidobacterium*, phylum Bacteroidales, order Bacteroidia, class Actinobacteria) were excluded as having a known biological origin of heterogeneity. **(B)** The correlation of taxon intercohort spread (calculated as $SD_{cohortMean}/E_{cohortMean}$) with the number mbQTLs detected per taxon with a relaxed threshold of $P<10^{-5}$. The genus *Bifidobacterium* and its related taxa of upper levels (family *Bifidobacterium*, phylum Bacteroidales, order Bacteroidia, class Actinobacteria) were excluded as having a known biological origin of heterogeneity. **(C)** The correlation of effective sample size (number of samples with non-zero bacterial abundance included in actual mbQTL analysis) with the number of detected mbQTL loci with at least one SNP with $P<10^{-5}$. Each point represents one taxon. A window of 1 mb was taken to define the locus**.**



5 **Figure S6.** Genome set enrichment analysis (GSEA) for mbQTLs. **(A)** GSEA analysis on 30 main tissue types. **(B)** GSEA analysis on 53 tissue types.

Supplementary tables legends

Table S1. Overview of cohorts
The table includes information on cohort size, ethnicity, details of microbial DNA isolation and sequencing and information on host genotyping, including genotypes and sample quality control.

Table S2. Participant descriptions per cohort
The table includes information on age, BMI and alpha diversity per cohort in all participants and also separately in males and females

Table S3. Bacterial abundance
The table indicates the abundance of all 257 bacteria included in quantitative or binary analyses. Columns QUANT and BINARY indicate if taxonomy was included in the corresponding analysis. Full.cohortsize: the total number of all participants in cohorts where this bacteria was analyzed. Full.nonzero - the number of all participants with non-zero abundance of this bacteria. Full.mean.presence - the proportion of bacteria in individuals who have non-zero abundance of this bacteria. Full.mean.Nreads - the mean number of reads in non-zero participants (out of 10,000).
Cohortsize.StoolMini;PowerSoil;NucleoSpin;DiaSorin;PowerLyzer;AllPrep;SpinStool;TissueLEV - number of samples isolated with corresponding stool DNA isolation kits.
Nonzeros.StoolMini;PowerSoil;NucleoSpin;DiaSorin;PowerLyzer;AllPrep;SpinStool;TissueLEV - number of non-zero samples isolated with corresponding stool DNA isolation kits.
MeanPresence.StoolMini;PowerSoil;NucleoSpin;DiaSorin;PowerLyzer;AllPrep;SpinStool;TissueLEV - mean presence of bacteria in the samples isolated by corresponding stool DNA isolation kits.
Nreads.StoolMini;PowerSoil;NucleoSpin;DiaSorin;PowerLyzer;AllPrep;SpinStool;TissueLEV - mean number of reads in samples isolated with corresponding stool DNA isolation kits.
Cohortsize.V12;V34;V4, Nonzeros.V12;V34;V4, MeanPresence.V12;V34;V4, Nreads.V12;V34;V4 - same parameters for samples that were sequenced using different 16S domains: V1-2, V2-3 and V4, respectively.

Table S4. Effect of DNA extraction method, domain selection and ethnicity

The estimation of effects of DNA extraction methods, temperature lysis, 16S domains selected for sequencing and cohort ethnicities on the microbial diversity (genus richness and beta diversity, see corresponding columns). The effects of technical covariates and ethnicity were on the level of summary statistics, where each cohort was represented as a vector composed of mean abundances of each genus/richness. The 'single model' represents univariate analysis of each of the technical parameters on the corresponding trait (linear regression for richness, dbRDA for beta diversity). The 'added model' represents the explained variance in stepwise forward model selection.

Table S5. Heritability analysis in twins in the TwinsUK and NTR cohorts, and by using LD score regression.

Heritability of the gut microbial taxa. The table contains the results of the heritability analysis. LD score regression heritability was performed of full summary statistics of mbQTLs. Heritability estimation (slope) with corresponding standard error and z-score are given in columns H2, SE, and Z respectively. The intercept of LD score regression and its standard error are given in columns Intercept and Int_SE columns. Twin-based heritability estiamation was conducted on 589 pairs of twins from the TwinsUK cohort using an ACE model (OpenMX in R). Estimates of A (genetic effects), C (common environment) and E (environment unique to an individual) are reported, as well as their 95% confidence intervals (CI_low/CI_high) and p-values (P). The number of monozygotic and dizygotic pairs for which the taxa were detected in both individual twins from a pair are indicated in the nMZ_pairs and nDZ_pairs, respectively. Finally, intraclass correlations between monozygotic twin pairs were calculated for each taxon in the TwinsUK (ICC_MZ) and NTR (ICC_MZ_NTR) cohorts. For the TwinsUK cohort, intraclass correlation was also calculated for dizygotic twin pairs (ICC_DZ).

Table S6. Effect of host genetics on bacterial alpha diversity

The table indicates the results of the analysis of the effect of SNPs on microbial alpha-diversity. SNPs with a p-value less than $1x10^{-5}$ are included. SNPname: Position in the genome is indicated in genome build NCBI37. TraitName: the alpha diversity metric for which association is reported (Shannon, Simpson and inverse-Simpson). Overall z-score: the z-score from the meta-analysis of all cohorts. DatasetsIncluded: datasets in which both the SNP and the microbial trait passed our

analysis threshold and were thus included in the analysis. DatasetsZScores: Z-scores per cohort. DatasetsNrSamples: number of samples per cohort in which both the SNP and trait were present.

Table S7. Binary and Quantitative microbiome trait loci (mbTLs) reached nominal genome-wide significance level ($P<5x10^{-8}$)

The table contains the associations of microbial traits with leadings SNPs for the 30 mbQTLs identified at the genome-wide significance (p<5x10-8) level. Only top SNP per locus is included. Type: the type of analysis (quantitative or binary). MAF_1000G: minor allele frequency of associated allele in 1000 genomes database. N_SNPs_$p<5x10^{-8}$: the number of SNPs in the locus that were associated with the corresponding bacteria at $p<5x10^{-8}$. Overall z-score: the Z-score from the meta-analysis of all cohorts. SampleSize: total number of samples included in the analysis. Genes in the locus: positional gene candidates. Cis-eQTLs: eQTL gene candidates. Trans-eQTLs: trans-eQTL gene candidates.

Table S8. Effect of host genetics on bacterial abundance

The table indicates the results of the analysis of the effect of SNPs on bacterial abundance. SNPs with a p-value less than $1x10^{-5}$ are included. SNPname. Position in the genome is indicated in genome build 37. TaxonName: taxonomy trait for which association is reported. Overall z-score: the z-score from the meta-analysis of all cohorts. DatasetsIncluded: datasets in which both the SNPs and the microbial trait passed cutoffs. DatasetsZScores: Z-scores per dataset. SampleSize: total number of samples included in the analysis. NCohorts: number of cohorts included in the analysis. DatasetsNrSamples: number of samples per dataset in which both the SNP and trait were present and included in the analysis. Nreads.mean: mean number of reads mapped to the taxon in the samples included in the analysis. Presence.mean: mean presence rate of the microbial trait among cohorts included in the corresponding mbQTL.

Table S9. Effect of host genetics on bacterial presence/absence

The table indicates the results of the analysis of the effect of SNPs on bacterial presence. SNPs with a p-value less than $1x10^{-5}$ are included. SNPname. Position in the genome is indicated in genome build 37. TaxonName: taxonomy trait for which association is reported. Overall z-score: the z-score from the meta-analysis of all cohorts. DatasetsIncluded: datasets in which both the SNPs and the microbial trait passed our cutoffs. DatasetsZScores: Z-scores per dataset. SampleSize: total number of samples included in the analysis. NCohorts: number of cohorts

included in the analysis. DatasetsNrSamples: number of samples per dataset in which both the SNP and trait were present and included in the analysis. Nreads.mean: mean number of reads mapped to the taxon in the samples included in the analysis. Presence.mean: mean presence rate of the microbial trait among cohorts included in corresponding mbBTL.

Table S10. Correlation of taxa with phenotypes in LLD cohort

The table contains the results of the correlation analysis of taxa that have at least one mbQTL or mbBTL at $P<5x10^{-8}$ and the 207 phenotypes measured in the LLD cohort as described in (Zhernakova et al, Science, 2016) [3]. MbTLtype: indication if the associated taxon has a mbQTL or mbBTL. Taxon: name of taxon. Phenotype: name of associated phenotype. SpearmanR: Spearman correlation between bacteria and phenotype. Pvalue: p-value for Spearman correlation. FDR: p-value adjusted using Benjamini-Hochberg procedure.

Table S11. Correlation of taxa with phenotypes in FGFP cohort

The table contains the results of the correlation analysis of taxa that have at least one quantitative or binary mbTL at $p<5x10^{-8}$ and the 503 phenotypes measured in the FGFP cohort as described in Falony et al, Science, 2016 [4]. Mb-TLtype: indication if taxon was associated in binary or quantitative analysis. Taxon: name of taxon. Phenotype: name of associated phenotype. SpearmanR: Spearman correlation between bacteria and phenotype. Pvalue: p-value for Spearman correlation. FDR: p-value corrected for false discovery rate.

Table S12. PheWAS analysis at single SNP level

Table shows the PheWAS lookups of the top SNPs in each genome-wide significant loci from the continuous trait GWAS of MiBioGen. In total 4155 traits were scanned using GWASATLAS database. Only result with P-value < 5 x 10-8 are shown. Sign adjusted Zscore MiBioGen indicates the Z-score for the from MiBioGen after shifting the effect allele to the same effect allele of the indicated PheWAS study. RG stands for genetic correlation, estimated by the LDSSCORE regression method. pRG is the significance of the correlation. PP.H4.abf, posterior probability of the SNP being causal SNP for both phenotypes, tested by the approximate Bayes factor approach. Genetic correlation analysis was only performed for *Bifidobacterium*, due to lack of heritability of other taxa.

Table S13. PheWAS analysis at gene level

The table contains the results of gene-based PheWAS lookups for mbTL. Atlas ID: identifier of phenotype in ATLAS database. PMID: PubMed ID of the manuscript providing summary statistics for the phenotype. Year: year of publication of GWAS study. Domain: the domain to which phenotype belongs. Trait: the conventional name of the phenotype. Pvalue: Pvalue of PheWAS signal. N study samples: number of samples included in the phenotype GWAS study. Gene name: name of the gene used in PheWAS analysis. Bacterial taxon: the taxon for which the mbQTL is reported.

Table S14. PheWAS enrichment analysis

The table contains the results of PheWAS enrichment analysis (see Methods). Each row represents the leading SNP of the mbQTL for which enrichment is reported. Each column represents the enrichment for the specific domain (odds ratio or p-value). The value in the corresponding cell shows the fold enrichment for the corresponding SNP-domain pair.

Table S15. Mendelian randomization analysis results at FDR < 0.05

The table contains a selection of Mendelian randomization (MR) results with a multiple testing adjusted p-value < 0.05 and that passed the major quality control steps. Results passing all quality control steps are highlighted in yellow. The multiple testing was corrected per category. We kept the results that failed our QC requirement that the nominal p-value of the MR test for the opposite direction (outcome -> exposure) was < 0.05 to show the cases with a complex causality structure. We also kept the results with a low heterogeneity test p-value if the MR PRESSO outlier test showed a significant result after outlier removal. These values that failed our QC are highlighted in red. Column description. exposure: MR exposure trait name; outcome: MR outcome trait name; pubmed_id: PubMed id of the source of the summary statistics; Nsnps number of instrumental variables; IVW_beta, IWV_se, IVW_pval: beta, standard error and p-value of the inverse variance weighted MR test; IVW_BH_qval: BH-adjusted p-value; SNPs: SNPs used as instrumental variables; egger_intercept_pval: p-value for the Egger test intersept; heterogeneity_Q_pval, heterogeneity_Q: heterogeneity test p-value and Q value; weighted_median_pval, weighted_median_beta, weighted_median_se: beta, standard error and p-value of the weighted median MR test;egger_pval, egger_beta, egger_se: beta, standard error and p-value of the MR_Egger test; mr_presso_pval, mr_presso_outlier_cor_pval,

mr_presso_global: p-values of the MR PRESSO test, MR_PRESSO global test, MR PRESSO test corrected for outliers; leave_one_out_pval: IVW p-values after rerunning MR with each SNP removed; reverse_direction_MR_pval: IVW MR p-value for the outcome -> exposure direction; F_statistics_exposure: F statistics for the exposure.

Table S16. Raw mendelian randomization analysis results before quality control

The table contains all Mendelian randomization (MR) results without any filters applied. Results with p-values adjusted for multiple testing < 0.05 and passed QC filters are highlighted in yellow. The multiple testing was corrected per category. Those values that failed QC are highlighted in red. Column description. Category: trait type used in MR; exposure: MR exposure trait name; outcome: MR outcome trait name; Nsnps: number of instrumental variables; IVW_beta, IWV_se, IVW_pval: beta, standard error and p-value of the inverse variance weighted MR test; IVW_BH_qval: BH-adjusted p-value; SNPs: SNPs used as instrumental variables; egger_intercept_pval: p-value for the Egger test intersept; heterogeneity_Q_pval, heterogeneity_Q: heterogeneity test p-value and Q value; weighted_median_pval, weighted_median_beta, weighted_median_se: beta, standard error and p-value of the weighted median MR test;egger_pval, egger_beta, egger_se: beta, standard error and p-value of the MR_Egger test; mr_presso_pval, mr_presso_outlier_cor_pval, mr_presso_global: p-values of the MR PRESSO test, MR_PRESSO global test, MR PRESSO test corrected for outliers; leave_one_out_pval: IVW p-values after rerunning MR with each SNP removed; reverse_direction_MR_pval: IVW MR p-value for the outcome -> exposure direction; F_statistics_exposure: F statistics for the exposure; final_filter: indication whether this MR result passed quality control; failed_filters: quality control steps that this MR result failed.-