



OPEN Synergy between culturomics and metagenomics of health status-associated gut bacteria originating from non-IBD and IBD populations

N. Plomp^{1,2}, R. Gacesa^{2,3}, J. Slager³, J. N. Samsom⁴, K. N. Faber², I. H. Jonkers³, S. Withoff³, C. Wijmenga³, R. K. Weersma² & H. J. M. Harmsen¹✉

The bacteria in the human intestinal tract are important for health and associate with diseases, such as inflammatory bowel disease (IBD). Although metagenomic studies can identify certain bacteria or even specific strains and associate their presence or specific phenotypes with health or diseases, actual isolates for experimental validation of metagenomic associations are often lacking. Therefore, this study sets out to culture health- and IBD-associated bacteria from 32 fecal samples from 2 cohorts, for which extensive metadata is available. The cultivation of those samples resulted in 4,347 isolates, of which 1,362 isolates were obtained from IBD patients. Irrespective of health or IBD, *Actinomyetota*, *Bacillota* and *Bacteroidota* were the most represented phyla and members of 5 other phyla were less frequently isolated (*Campylobacterota*, *Fusobacteriota*, *Pseudomonadota*, *Thermodesulfobacteriota* and *Verrucomicrobiota*). Comparison of the genus richness between the culturomics approach and available metagenomic sequencing data of the corresponding participants revealed that both methods largely capture the same genera. Although not all genera could be identified in both methods, our results show that combining both methods has a synergetic effect, providing a higher identification rate. Furthermore, genetic analysis of 2 isolates of *Bifidobacterium adolescentis* strains shows that these isolates closely resembled the metagenome-assembled genome that was identified within the same participant. This showcases that it is possible to isolate specific strains that are important in the experimental validation of specific associations within a species. The culture collection that is presented in this study contains bacterial isolates that are strongly associated with health or IBD. Our results show that we are able to generate a valuable culture collection that opens a promising avenue for functional validation experiments of associations that are identified with metagenomic data.

Keywords Human gut microbiota, Metagenomics, Culturomics, Inflammatory bowel disease, Serial dilution

The human gut microbiota plays a crucial role in maintaining the health and metabolism of the host^{1–3}. This diverse gut ecosystem is composed of a complex array of microorganisms, including bacteria, archaea, protozoa, and eukaryotes. Yet, despite this vast richness of over 5,600 gut bacterial species discovered so far^{4,5}, individuals typically harbor a core set of approximately 160 bacterial species⁶, which predominantly belong to 3 major phyla: the *Bacillota*, *Bacteroidota* and the *Actinomyetota* (formerly known as *Firmicutes*, *Bacteroidetes* and *Actinobacteria* respectively)^{6,7}. This vast collection of diverse bacterial species is well-balanced in healthy individuals and acts as an extension of the host metabolism and gene repertoire and therefore has an important role in the host physiology, for instance by producing essential amino acids, short chain fatty acids, vitamins, secondary metabolites and hormones^{8–10}. The gut microbiota also regulates and interacts with the immune system^{11,12} and protects the host from colonization of foreign pathogens, for instance with short-chain fatty acid production¹².

¹Department of Medical Microbiology and Infection Prevention, University of Groningen, University Medical Center Groningen, Hanzeplein 1, 9713 GZ Groningen, The Netherlands. ²Department of Gastroenterology and Hepatology, University of Groningen, University Medical Center Groningen, Hanzeplein 1, 9713 GZ Groningen, The Netherlands. ³Department of Genetics, University of Groningen, University Medical Center Groningen, Hanzeplein 1, 9713 GZ Groningen, The Netherlands. ⁴Laboratory of Pediatrics, Division of Gastroenterology, Erasmus MC Sophia Children's Hospital, Dr. Molewaterplein 40, 3015 GD Rotterdam, The Netherlands. ✉email: h.j.m.harmsen@umcg.nl

An imbalance, defined as dysbiosis, of the gut microbiota is a key feature of inflammatory diseases, such as inflammatory bowel disease (IBD), a group of chronic intestinal inflammatory disorders that affect millions of people worldwide¹³. A reduced microbial richness, especially of strict anaerobic bacteria, and a shift in relative abundance of the remaining bacterial species are the most consistent observations of gut microbial dysbiosis in IBD. The latter observation is often characterized by a decreased relative abundance of *Bacillota* species, whereas the relative abundance of *Pseudomonadota* (formerly known as *Proteobacteria*) species is increased^{14,15}. These compositional changes contribute to altered metabolic functions of the gut microbiota in IBD, where IBD patients display decreased short chain fatty acids^{16,17} and secondary bile acids levels^{18,19}. Recent work has shown that the compositional changes of the gut microbiota are not limited to phylum or even species level, but that the presence of specific strains of individual species are associated with the health status of the host and these specific strains outcompete other strains during inflammation, which was described in detail for *Eggerthella lenta*²⁰. Reconstruction of the genetic differences between health- and disease-associated strains identified simultaneous changes in various disease-associated pathways, including oxidative stress and nutrient biosynthesis²⁰. These changes in multiple pathways provide an explanation for the relative fitness advantage of particular disease-associated strains and demonstrates that the entire disease environment selects for certain strains, rather than one specific factor. While these findings represent a significant advancement in understanding the ecological changes in the IBD microbiome, experimental validation is still required to validate the hypothesis of fitness advantage mediated by particular pathway adaptations and possible strain-to-strain interactions of certain strains that are identified based on metagenome-assembled genomes (MAGs).

Bacterial culture collections that contain health- and disease-associated isolates are indispensable for functional validation of associations and permit complementary research into the interaction of different strains or the interaction of certain strains with the host. The inclusion of samples from hosts with diverse health statuses, such as IBD patients, into culturomics efforts is therefore essential. Several efforts focused on healthy subjects have expanded the human gut culturome^{21–26}, but there is a lack of culture collections targeting IBD patients. While previous efforts relied on serial dilution of samples before inoculating the media to obtain single colonies, a recent study advocate omitting the dilution step and directly inoculating the samples onto the media²⁷. This approach is thought to better preserve the bacterial community, potentially facilitating in vivo-like bacterial interactions, stimulating the growth of fastidious bacteria, and increasing the species richness of the isolates in the primary inoculation of the media.

The recent culturomics effort proposed a convenient protocol using a single basal medium, primarily focused on isolating *Bacillota* species, particularly from the *Lachnospiraceae* and *Oscillospiraceae* families²⁷. Building upon this platform, the current study aims to capture a broader diversity of phyla and families using 16 anaerobic growth media and 4 aerobic growth media, with an emphasis on IBD-associated strains that were previously identified using metagenomic sequencing¹⁵. This study also performed targeted isolation of IBD-associated strains reconstructed from metagenomics data to optimize the standard culturomics protocol, generating a comprehensive collection of gut bacteria to facilitate functional studies of strain-strain and strain-host interactions in IBD.

Methods

Sample collection

Fecal samples were collected from volunteers of 2 different study populations, namely, the Lifelines Dutch Microbiome Project cohort¹⁵ and the RISE-UP study cohort²⁸. For the RISE-UP study, patients with an established diagnosis for at least 1 year were included from March 2016 until April 2017 from the IBD outpatient clinic of the University Medical Center Groningen. The diagnosis was based on clinical, histopathological and endoscopic criteria according to the diagnostic guidelines²⁹. Patients were excluded from the RISE-UP study sample collection if any of the following criteria was met: use of antibiotics, probiotics or prebiotic supplements in the 3 weeks prior to collection; pregnancy or lactation during the collection period; colonoscopy or colon cleansing in the 3 months prior to the collection; use of a vitamin B2 supplement or other multivitamin complexes that contain B vitamins in the 3 weeks before collection; patients with severe CD activity with a Harvey-Bradshaw Index higher than 12; methotrexate drug usage. Four fecal samples of IBD patients that originated from the initial collection time point were chosen for further bacterial isolation. The RISE-UP study was approved by the medical ethical committee of the University Medical Center Groningen (02/09/2015; METc numbers 2008/338, 2014/291 and 2016/424) and was registered on ClinicalTrials.gov (NCT02538354). In the Lifelines Dutch Microbiome Project (DMP), fecal samples of Lifelines DMP participants were collected between 2015 and 2016. Metadata was collected using questionnaires and the data was curated as described previously^{15,30}. The Lifelines study was approved by the medical ethical committee of the University Medical Center Groningen under the METc number 2017/152. All research involving human participants have been performed in accordance with both the Declaration of Helsinki and local guidelines and regulations. For our study, 15 random fecal samples of non-IBD participants of the Lifelines DMP and 17 random fecal samples of IBD patients (13 Lifelines DMP and 4 RISE-UP study) were selected based on the metadata. Participants with antibiotic usage within 3 months prior to the collection data were excluded from this study. Basic information, such as sex, age, and health status are summarized in Supplementary Table 1.

Bacterial cultivation procedures

The fecal samples were collected in 2 ml eSwab® Amies medium that was supplemented with 20% glycerol (Copan Group, Brescia, Italy). The samples were immediately frozen after collection and stored at -80°C until further analysis. The samples were thawed and homogenized and subsequently inoculated inside an anaerobic chamber for the anaerobic media and an aliquot of the sample was used for inoculation at atmospheric conditions for the aerobic media. The inoculation with the 'direct inoculation method' was performed with 50 μl of homogenized

sample that was directly spread onto the media in 3 segments without preceding dilution steps, as described before²⁷. An exception was an ethanol shock treatment condition, for which 25 µl of homogenized sample treated with a 1:1 ratio of absolute ethanol for 30 min and subsequently inoculated on brain heart infusion agar (BHI; Oxoid Limited, Hampshire, UK). The 'dilution method' for the anaerobic condition comprised 10^{-4} , 10^{-5} , 10^{-6} and 10^{-7} dilutions of the original fecal sample in overnight reduced phosphate buffered saline (PBS), whereas 10^{-2} , 10^{-3} , 10^{-4} and 10^{-5} dilutions were made in non-reduced PBS for the aerobic condition. The anaerobic media were incubated in an anaerobic chamber (Whitley A35 Workstation, Don Whitley Scientific Limited, West Yorkshire, UK) with an anaerobic gas mixture (10% H₂, 10% CO₂, and 80% N₂) and the aerobic media were incubated in an atmospheric incubator. All samples were incubated at 37 °C for 48 h before picking colonies. For every plate, all morphologically diverse colonies were selected and streaked onto fresh agar until single colonies were obtained. The preliminary purity of the single colonies was assessed using Gram staining, after which the identity was determined with matrix-assisted laser desorption/ionization coupled to a time-of-flight mass spectrometer, as described below.

Media for bacterial cultivation

For this study, various growth conditions and cultivation media were used in order to isolate a broad richness of bacteria within the samples. Different media were used for aerobic and anaerobic incubation of the samples. For the aerobic media with the 'direct inoculation method', blood agar (BA; Media Products BV, Groningen, The Netherlands); esculin azide agar (Media Products BV); MacConkey agar (no. 3; Media Products BV) and mannitol salt agar (Media Products BV) were used. The aerobic 'serial dilution method' was performed on BA (Media Products BV). The following anaerobic media were used for the 'direct inoculation method': Brucella blood agar (BBA; Media Products BV); *Fusobacterium*-specific medium (FSM; Media Products BV); phenylethyl alcohol agar (PEA; Media Products BV); brain heart infusion agar supplemented with 1 g/l cysteine, with and without ethanol shock treatment; De Man Rogosa Sharpe agar (MRS; Oxoid Limited, Hampshire, UK); minimal medium supplemented with melezitose; YCFA supplemented with glucose (YCFAG); YCFA supplemented with glucose and 20 µg of sulfamethoxazole (YCFASMX); YCFA supplemented with glucose and cellobiose (YCFAGC); YCFA supplemented with kiwi and apple pectin (YCFAKAP); YCFA supplemented with porcine mucin type III (YCFAM); YCFA supplemented with melezitose (YCFAMel); YCFA supplemented with apple pectin (YCFAP); YCFA supplemented with xylose (YCFAX) and YCFA supplemented with inulin (Orafti inulin, Beneo GmbH, Mannheim, Germany), xylose and fructose (YCFAIXF). The anaerobic 'direct inoculation method' was performed on BBA (Media Products BV), YCFAGC and YCFAP. The supplemented carbohydrates are added in a concentration of 4.5 g/L, as previously described²⁷, and combinations of carbohydrates were added in equal amounts (w/w) to an end volume of 4.5 g/L. The minimal medium with melezitose and all YCFA-based media were prepared in-house. Detailed information on the composition of the media that were prepared in-house, and the selectivity of all media are found in Supplementary File 1 and 2.

Identification of the isolates

The identity of the pure bacterial isolates was determined with MALDI-TOF MS (Biotyper Microflex, Bruker Daltonics, Billerica, USA). For this, a single colony of each isolate was spotted in duplicate onto a polished stainless steel MALDI target. One of the duplicates was covered with 1.0 µl of 70% formic acid and was allowed to air-dry. Next, all duplicate spots were covered with 1 µl of 2-Cyano-3-(4-hydroxyphenyl) acrylic acid matrix (10 mg/ml) and were air-dried again. Subsequently, spectra were obtained from these spots by summing 240 measured shots (6 times 40 shots each) of the Biotyper Microflex. The laser was set to have a minimum power of 30% and a maximum power of 40%³¹. The summed spectra were analyzed using the commercially available database (Bruker Daltonics; database version 11) with clinically relevant bacteria and subsequently using the publicly available database 'MaldiGut' that contains anaerobic gut bacteria²⁷. An isolate was considered to have a reliable identification at the species level in case of a detection score with a log of 2.0 or higher. If the log of the score was between 1.7 and 2.0, the isolate was considered to have a reliable identification to the genus level. A log smaller than 1.7 indicated no reliable identification and was therefore considered as 'no ID'.

16S rRNA gene Sanger sequencing

A selection of bacteria that were suspected to be species that are highly associated with health status¹⁵ or hard to distinguish between closely related species based on MALDI-TOF MS were subsequently analyzed with 16S rRNA gene sequencing. A full 1.0 µl loop of bacteria from 2 or more single colonies, derived from a pure culture, was dissolved in 50 µL lysis buffer, containing 0.05 mM NaOH and 0.25% SDS in Milli-Q water (Sigma-Aldrich, St. Louis, Missouri, USA). The lysis buffer was incubated at 100 °C for 10 min and afterwards 200 µl of 1 × TE buffer (10 mM Tris-HCl and 1 mM EDTA, pH 8.0) was added to the lysate. The primary amplification was performed using a PCR with 0.2 µM of the universal primers 27F and 1492R or 515R³², 50% v/v 2X Phire Hot Start II PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA) and 37.5 ng DNA dissolved in nuclease-free water, which resulted in an end volume of 25 µl. The PCR program consisted of one denaturation step at 98 °C for 30 s, followed by 35 cycles of denaturation (98 °C for 10 s), annealing (60 °C for 10 s) and extension (72 °C for 15 s), ending after these cycles with a final step of incubation at 72 °C for 2 min. The amplicons were sequenced on the ABI-3500XL genetic analyzer with the BigDye Terminator v3.1 Cycle Sequencing kit according to the manufacturer's instructions (Applied Biosystems, Waltham, Massachusetts, USA). Curation of the reads was performed manually in BioEdit v5.0.9, in which also the consensus sequences were calculated. The consensus sequences were annotated to genus level using the Ribosomal Database project (RDP; release 11) in the RDP classifier tool and EzBioCloud was used to annotate these consensus sequences to the species level^{33–35}.

DNA extraction and whole genome sequencing of isolates

The genomic DNA of pure isolates was extracted with the Ultraclean Microbial DNA Isolation Kit (QIAGEN, Hilden, Germany). The extraction was performed with a full 10 µl loop of bacteria from 2 or more single pure colonies, further steps were done according to the protocol. The DNA quality was assessed with a 2000 NanoDrop UV-Vis spectrophotometer (Thermo Fisher Scientific) by determining the A_{260}/A_{230} and A_{260}/A_{280} ratios. An A_{260}/A_{230} ratio of 2.0 to 2.2 and an A_{260}/A_{280} ratio between 1.7 and 2.0 were considered as pure nucleic acid extracts. The DNA concentrations were determined with the dsDNA HS or dsDNA BR assay kit using the Qubit 2.0 fluorometer (Life Technologies, Carlsbad, CA, USA). The extracts were used for DNA library preparation using the Nextera XT v2 Library kit (Illumina, San Diego, CA, USA) according to the protocol. The DNA libraries were subsequently used for short-read sequencing on the in-house Illumina MiSeq platform that resulted in paired-end reads of 250 bp. De novo assembly of these paired-end reads was performed in the CLC Genomics Workbench v12.0.1 (Qiagen, Hilden, Germany), using a word size of 29 after quality trimming ($Q_s \geq 20$).

DNA extraction and metagenomic sequencing of fecal samples

The microbiome composition of the fecal samples was characterized with metagenomic shotgun sequencing. For the RISE-UP study samples, the microbial DNA was extracted with the Qiagen Allprep DNA/RNA Mini Kit and the Nextera XT v2 Library kit (Illumina) was used for the subsequent library preparation. The metagenomic shotgun sequencing was executed at the Broad Institute of the Harvard University and the Massachusetts Institute of Technology (Cambridge, MA, USA) on the Illumina HiSeq 2000 platform. For the Lifelines DMP cohort, the microbial DNA extraction was performed with the QIAamp Fast DNA Stool Mini Kit (Qiagen), following the protocol of the manufacturer, with the QIAcube automated sample preparation system (Qiagen). Afterwards, the DNA concentration was determined with the Qubit 4 fluorometer (Life Technologies). The library preparation was done with either the NEBNext Ultra DNA Library Prep Kit for Illumina if the DNA yield was lower than 200 ng, or with the NEBNext Ultra II DNA Library Prep Kit for Illumina if the yield was higher. The metagenomic shotgun sequencing was performed at Novogene China on the Illumina HiSeq 2000 platform.

Microbiome profiling of the fecal samples

Raw metagenomic reads of the RISE-UP study were cleaned by removing the sequencing adapters and trimming the ends with Trimmomatic (version 0.32)³⁶. The raw reads of each sample from the Lifelines DMP cohort were cleaned of sequencing adapters with the BBduk tool from the BBMap package (version 38.93) and subsequently trimmed and decontaminated using the KneadData toolkit (version 0.5.1), that integrates the Trimmomatic (version 0.39.2) and the Bowtie2 tool (version 2.3.4.1)³⁷. Trimmomatic was used for the removal of the sequencing adapters and the trimming of the reads to PHRED quality 30. The BowTie2 tool removed the reads that aligned to the reference human genome GRCh37/hg19. The remaining reads were used for the profiling of the bacterial taxonomic composition with the MetaPhlAn2 tool (version 2.7.2)³⁸. Sequencing generated 8 Gb of 150 nt paired-end reads per sample (mean = 7.9 Gb, s.d. = 1.2 Gb). Read quality was: Q1: 38; median, and Q3 of 40, with 0.26% of bases below Q20 and 4.7% below Q30. The MAGs were extracted from the clean reads by assembly with the MEGAHIT assembler (version 3.14.1) followed by binning and refining with the MetaWRAP pipeline (version 1.3.2), that uses the MaxBin2 (version 2.1.1), MetaBat2 (version 2.12.1) and the Concoct (version 1.0.0) algorithms for binning. The quality of the MAGs was evaluated with the CheckM tool (version 1.0.12), which assigned completeness and the potential contamination rate to the individual MAGs. The MAG identification, the completeness, and the contamination are summarized in Supplementary Table 2. The complete MAG binning pipeline is also published on https://github.com/GRONINGEN-MICROBIOME-CENTRE/GMH_MGS_pipeline.

Digital DNA-DNA hybridization and single nucleotide polymorphism analysis

The MAGs of *Bifidobacterium adolescentis* from participants from which this bacterial species could be isolated were used for the calculation of the digital DNA-DNA hybridization (dDDH) and average nucleotide identity (ANI) percentage based on the genome sequences of the isolates. The web-based Genome-to-Genome distance Calculator was used with BLAST + as local alignment tool (<https://ggdc.dsmz.de/ggdc.php#>) for the calculation of the dDDH score and the ANI values were determined with the ANI calculator that uses the OrthoANIu algorithm³⁹. The MAGs that had a dDDH score $\geq 99.5\%$ and an ANI score $\geq 99.90\%$ were used for subsequent alignment of the genome sequences of the isolates from the corresponding participant. The subset of MAGs were subsequently used as reference sequences to align the sequences of the associated isolates to perform single nucleotide polymorphism (SNP) calling and SNP filtering with a minimal SNP depth of 10 times, a relative depth of 10%, a minimal distance of 10 bp, a minimal quality score of 30, a minimal read mapping score of 25 and a minimal Z-score of 1.96. The SNP calling was performed on a genome alignment of the MAG 7010000598825.5, strain 8825-Y2b and 8825-B1 that covered 1,980,540 nucleotide positions and a genome alignment of MAG 7010000945627.30 and strain 5627-YX7b with 1,320,098 covered nucleotide positions. After SNP calling and filtering, an alignment was created with the concatenated nucleotides of the respective SNP positions.

Identification of the core genome and accessory genes in MAGs and isolates

The gene prediction and annotation were performed with Prokka (version 1.14.6) using the default settings with 0.000001 as similarity e-value cut-off and enabling the search for non-coding RNA with Infernal and Rfam⁴⁰. The ROARY tool (version 3.11.2) was used to calculate the pan genome with a minimum of 95% identity for BLASTp protein prediction and the requirement that the genes had to be present in 100% of the isolates to be considered core genes⁴¹. The evolutionary distances between the alignments of the core genomes was calculated with the maximum likelihood method (Tamura-Nei model)^{42,43} based on the nucleotide substitution rate of all

positions excluding gaps or missing data. Afterwards, the phylogenetic tree was constructed with the MEGA software (version 11)⁴⁴.

Statistical analysis

The statistical analysis was performed with R 4.4.1, determining the statistical significance with a paired t-test. The differences were considered significant if the p-value was ≤ 0.05 . The normality of the data was tested with the Shapiro–Wilk's test, assuming normality with p-values > 0.05 .

Results

The culturomics of fecal samples with diverse conditions

The implementation of protocols to culture and isolate a large phylogenetic richness of anaerobic and aerobic bacteria from 32 fecal samples resulted in a total of 4,347 pure isolates, of which 1,362 were obtained from IBD patients and 2,985 from non-IBD volunteers (Fig. 1). The protocols were based on a previous culturomics effort that was established in-house²⁷ and was built upon with different media, antibiotics, ethanol shock and oxygen requirement. The isolation resulted in at least 201 different identified species and 169 isolates that were not identified to species-level based on the current MALDI-TOF MS databases used in this study. Most bacterial isolates (3,552) were retrieved from anaerobic culture media. The anaerobic isolates belonged to 8 phyla, represented by one or more species (the number of different identified species per phylum is indicated between brackets). The 3 dominant phyla were *Actinomycetota* (25), *Bacillota* (103) and *Bacteroidota* (30). The other 5 phyla that were less frequently isolated were *Campylobacterota* (1), *Fusobacteriota* (3), *Pseudomonadota* (5), *Thermodesulfobacteriota* (1), and *Verrucomicrobiota* (1). Another 145 anaerobic isolates could not be identified based on the available MALDI-TOF MS databases and therefore these isolates might include species that are uniquely identified based on metagenomics in this study. A total of 795 pure colonies were isolated from aerobic culture media, of which 781 were successfully identified as *Bacillota* (46), *Pseudomonadota* (9) and *Actinomycetota* (11). The other 14 aerobic isolates could not be identified based on the available MALDI-TOF MS databases. In total, 8 phyla, 14 classes, 24 orders, 39 families and 84 genera were represented in our identified isolates (Supplementary Table 3).

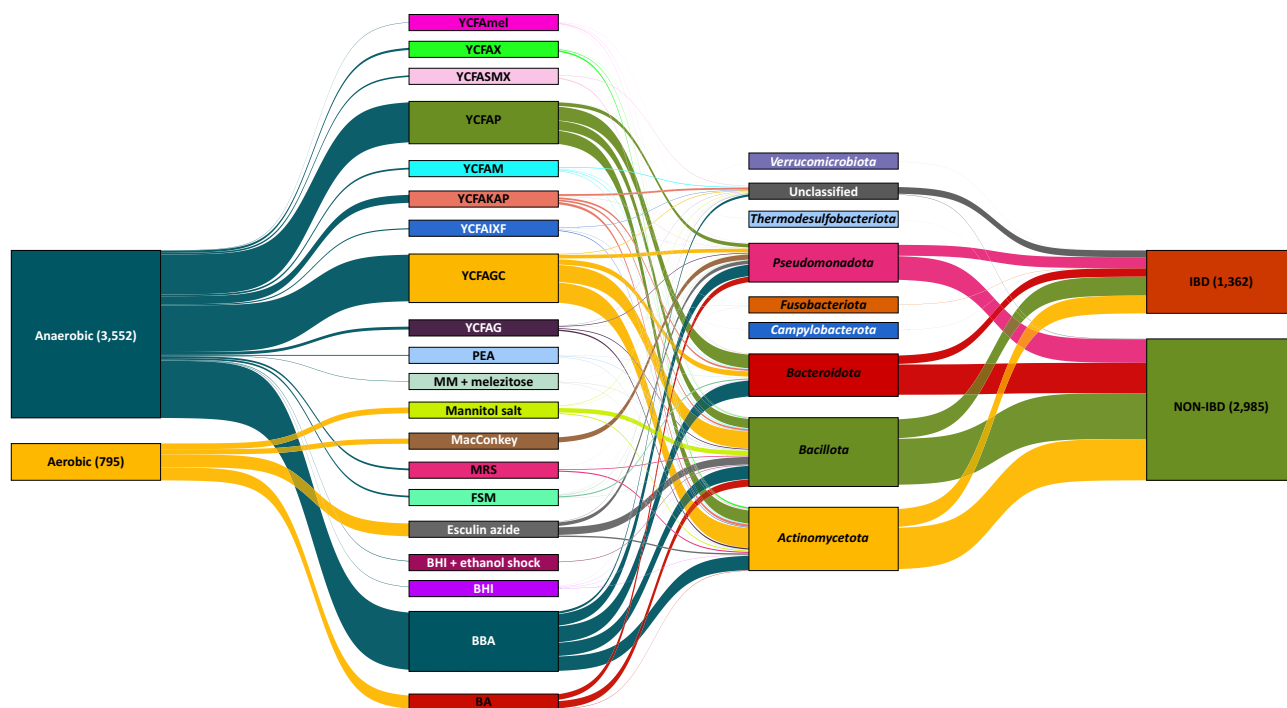


Fig. 1. Characteristics of the bacterial isolates. Sankey diagram of the 4,347 isolates obtained in this study, which were isolated from 17 fecal samples of IBD patients and 15 fecal samples of non-IBD participants. Each isolate is represented by a line that indicates the culture condition, isolation medium, the phylum to which each isolate belongs and the health status of the participant from whom the fecal sample was obtained (from left to right). Each path flows from left to right and is color-coded corresponding to the condition it originates. The bacterial isolates represent eight identified phyla: *Actinomycetota* (1250 isolates), *Bacillota* (1387 isolates), *Bacteroidota* (815 isolates), *Campylobacterota* (1 isolate), *Fusobacteriota* (4 isolates), *Pseudomonadota* (728 isolates), *Thermodesulfobacteriota* (1 isolate), and *Verrucomicrobiota* (1 isolate). The highest numbers of isolates were obtained anaerobically on YCFAP, YCFAGC or BBA medium. The cultures of fecal samples from IBD patients resulted in 1,362 isolates and all others were obtained from fecal samples of non-IBD participants.

Direct inoculation of fecal samples yields a higher species richness of anaerobic *Bacillota* and *Pseudomonadota* as compared to serial dilutions

The species richness of the isolated bacteria was compared between 2 different inoculation methods. A previous culturomics effort suggested that omission of dilution steps prior to inoculation might result in a higher species richness, especially of *Bacillota* species, compared to methods that use prior dilution steps²⁷. Therefore, a comparison of the species richness was performed onto BBA, YCFAGC and YCFAP media in anaerobic conditions with 12 individual samples. The direct inoculation of the samples resulted in a significantly higher species richness of the identified isolates that belonged to the *Bacillota* (Fig. 2a; p-value = 0.00077) and *Pseudomonadota* (Fig. 2b; p-value = 0.02097), as compared to the serial dilution method. On average, the direct inoculation method yielded more than twice the number of different species that belong to the *Bacillota* and 1.5 times the number of different species that belong to the *Pseudomonadota*. On the other hand, the serial dilution method resulted in a higher isolated species richness of the *Actinomycetota* (Fig. 2c; p-value = 0.00772) with on average 1.3 times more richness of the isolated species. The isolated species richness within the *Bacteroidota* did not significantly differ between the inoculation methods (Fig. 2d).

Culturomics and metagenomic assembled genomes largely capture the same genera

The overlap of the genus-level richness of the isolated culture collection was compared with the richness that was captured with a culture-independent method, based on extracted MAGs of the fecal samples that were collected from the same participants. Genus-level instead of species-level comparison was performed to make the analysis more robust, since not all isolates or assemblies were confidently identified to species-level. Genera that were included in the employed MALDI-TOF MS databases were exclusively evaluated for this analysis, since in this study, solely these genera were identified with the culture-based approach. The comparative analysis identified, 93 genera in total, either captured with the culturomics approach or with metagenomic sequencing (Fig. 3a). The culture collection contained 84 different identified genera, which is 90.3% of the identified genus richness. Sixty-one different genera were captured by either of the approaches, representing 65.6% of the genus richness. The culture-based approach did not retrieve isolates of the 9 genera *Bilophila*, *Coprobacter*, *Dialister*, *Eubacterium*, *Lachnospira*, *Megaphaera*, *Paraprevotella*, *Parasutterella*, and *Prevotella* that were captured within the MAGS, representing 9.7% of the identified richness. Interestingly, 23 different identified genera were unique to the culture-based approach, as metagenomic sequencing was unable to detect these genera that represented 24.7% of the genus-level richness. Analysis of the genus richness within each phylum as detected based on metagenomic sequencing (Fig. 3b), showed that the culture approach captured all genera that belong to the *Actinomycetota* (17 genera), *Campylobacterota* (1 genus; *Campylobacter*), *Fusobacteriota* (1 genus; *Fusobacterium*) and *Verrucomicrobiota* (1 genus; *Akkermansia*). In addition, most of the genera that belong to *Bacillota* and *Pseudomonadota* were represented in the culture collection, since the weighted percentage of captured genera that belong to the *Bacillota* was 91.8% and 90.0% for the genera that belong to the *Pseudomonadota*. Most of the genera within the *Bacteroidota* phylum (75%; 9 out of 12 genera) and half of the genera belonging to *Thermodesulfobacteriota* (1 genus; *Desulfovibrio*) were also isolated.

The culturomics approach yielded bacterial species that were associated with the host health status

Associations of certain bacterial species with the health status of the respective hosts were previously determined based on the Lifelines DMP cohort¹⁵. In that study, the top 40 different bacterial species were identified as being most strongly associated with the host health status (Fig. 4; Y-axis). The relative abundance of 16 out of these 40 species were positively associated with health of the host, whereas the relative abundance of 24 species were associated with IBD in the host. Bacterial cultures of the bacterial species that are strongly associated with health or IBD are needed to study the mechanistic role and causality within the host. Therefore, MAGs of species that are strongly associated with health and IBD were extracted within the samples that were used for culturing and the ability to culture these bacterial species from the corresponding samples was subsequently evaluated (Fig. 4). The presence of reference spectra within the employed MALDI-TOF MS databases was also taken into consideration to indicate which isolated bacteria could be identified in this study. Figure 4 shows that reference spectra of 31 species that were strongly associated with health or IBD were included in the databases. Twenty-two of these bacteria were retrieved from the samples, of which 20 were bacteria that were identified to species level based on extracted MAGs (Fig. 4). Eighteen of the bacterial species were also detected with the culture-independent methods. Interestingly, the MAGs of *Bifidobacterium dentium* and *Hungatella hathewayi* were not detected based on the current sequencing depth, yet isolates were obtained. MAGs of 11 other bacteria of interest were detected that represented 7 bacterial species, of which 4 species were included in the MALDI-TOF MS databases (Fig. 4). These results show that most of the 40 bacteria that are strongly associated with health or IBD that were identified in the Lifelines DMP cohort could be isolated from the respective sample population.

Specific metagenome-assembled genomes can be isolated with fecal culturomics

Since the bacterial species that had strong associations with host health status could be isolated from the corresponding sample population, it was investigated whether it was possible to isolate specific bacterial strains that are identical to MAGs that were extracted from the metagenomic data of the corresponding sample. The whole genome was determined of 3 *Bifidobacterium adolescentis* isolates from participants in which this species was also detected with metagenomic sequencing. Two isolates were isolated from different media that were inoculated with the same fecal sample (8825-Y2b and 8825-B1) and 1 remaining isolate (5627-YX7b) was obtained from a fecal sample of another participant. Digital DNA-DNA hybridization (dDDH) and calculation of the average nucleotide identity were performed as primary tools to select strains that are closely related (dDDH score $\geq 99.5\%$ and ANI score $\geq 99.90\%$). Both strains 8825-Y2b and 8825-B1 had a dDDH scores of

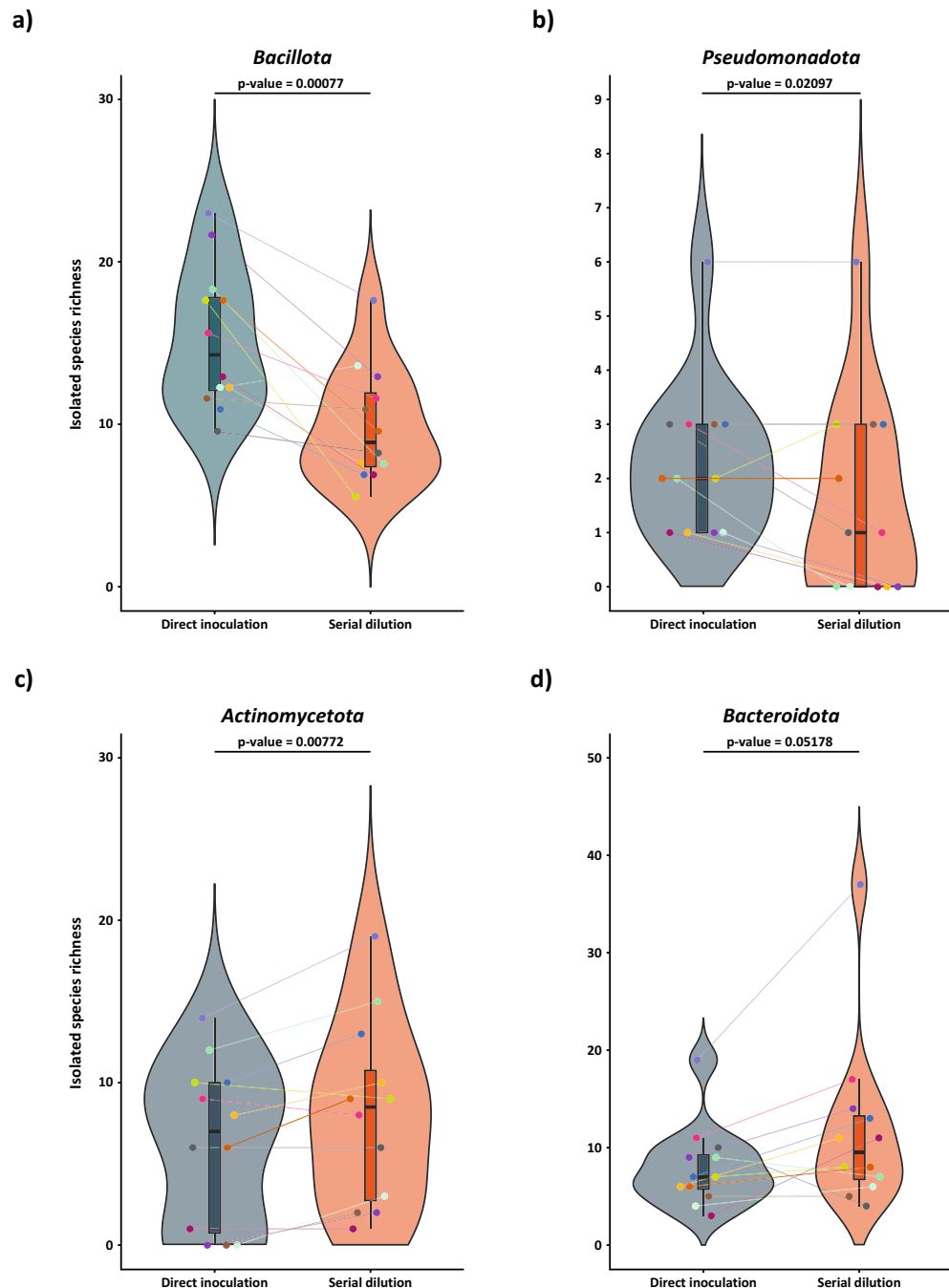


Fig. 2. Impact of the inoculation method on the isolated bacterial richness. Violin plots of the isolated species richness of anaerobic *Bacillota* (a), *Pseudomonadota* (b), *Actinomycetota* (c) and *Bacteroidota* (d). The species richness of the 4 phyla that were isolated with the direct inoculation method (blue) was compared to the serial dilution method (red). The individual values per sample are displayed as dots and the values are paired by lines. The direct inoculation method resulted in a higher isolated species richness of anaerobic *Bacillota* and *Pseudomonadota* and the isolated species richness of *Actinomycetota* was significantly higher with serial dilution. The difference between the 2 inoculation methods was not significant for *Bacteroidota* species. The paired t-test was performed to calculate significance and adjusted p-values ≤ 0.05 (Bonferroni correction) were considered significant.

99.9% and ANI values of 99.97% MAG 7010000598825.5, indicating that the strains closely resemble the MAG of the corresponding participant. In addition, the 2 strains were also closely related to each other (dDDH score: 99.9%; ANI score: 99.97%). The *B. adolescentis* strain 5627-YX7b was isolated from another participant and had a dDDH score of 99.7% and an ANI score of 99.92% with the corresponding MAG 7010000945627.30. A single nucleotide polymorphism (SNP) analysis was subsequently performed on the 3 strains to determine the strain-relatedness of the isolates and the related MAGs. The SNP analysis based on the genome alignment with the

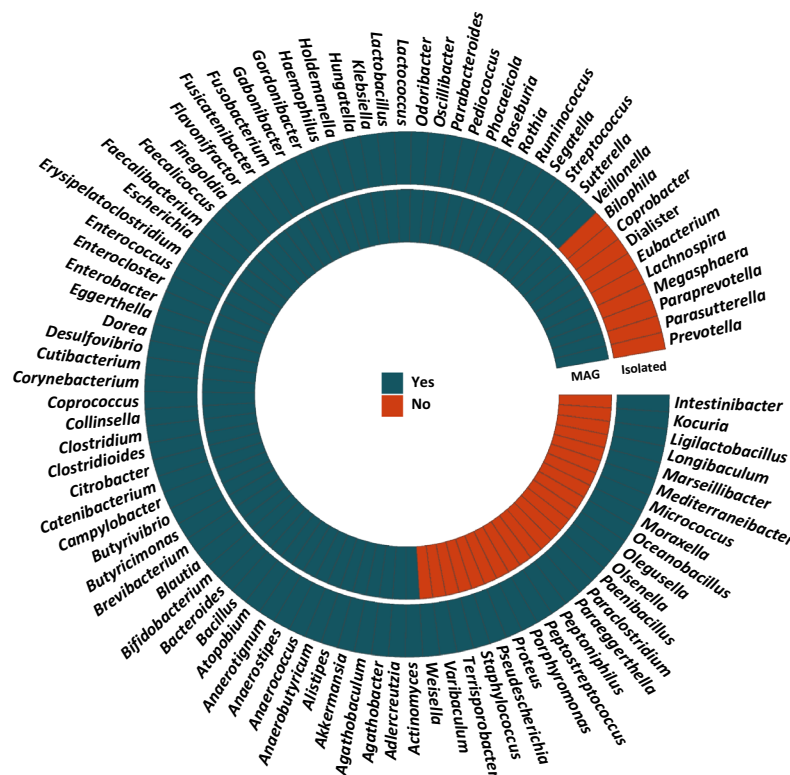
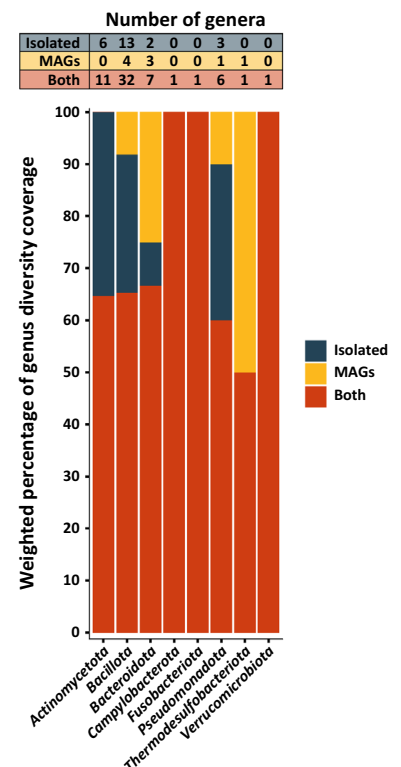
a**b**

Fig. 3. Comparison of culture-dependent and culture-independent genus diversity **a**) Circular heat map representation of the genera that were detected as MAG with metagenomic sequencing (inner ring) or that could be isolated in this study (outer ring). The presence (yes; blue) or absence (no; red) of the individual genus is indicated for each method. Genera that were included in the MALDI-TOF MS databases were exclusively selected for comparison, which shows that 60 genera were captured in both methods, 9 genera were uniquely found as MAG, and 24 genera were isolated, but not captured during metagenomic sequencing. **b**) The phylum distribution of the genus comparison between the culture and culture-independent method. The weighted percentage and number of genera are indicated for the isolated genera in blue (Isolated), for the genera that were found as MAGs in red (MAGs) and the genera that were found in both methods in green (Both).

MAG of the isolate that was extracted from the corresponding participant as reference genome confirmed that the strains 8825-Y2b and 8825-B1 closely resemble MAG 7010000598825.5 with only 20 and 21 SNPs difference based on the covered positions that were evenly distributed across the reference genome sequence and did not cluster in a specific region (Fig. 5a), which is within the limits of what we consider the same strain⁵⁸. The strain 5627-YX7b differed 134 SNPs with MAG 7010000945627.30. Next, gene annotation was performed based on the 3 genome sequences of MAG 7010000598825.5, 8825-Y2b and 8825-B1, to determine theoretical functional differences despite the limited SNP differences. This revealed that both *B. adolescentis* strains and the MAG from participant 7010000598825.5 share a core genome of 1,735,932 nucleotides that encoded 1,548 different genes, where both isolates were only missing 21 genes that were found in the MAG 7010000598825.5 (Fig. 5b), encoding for hypothetical proteins. The 21 genes of the MAG that were missing in the isolates were investigated to determine the potential relevance of the lack of these genes. Closer inspection of the gene annotations revealed that although both strains were missing 21 genes as compared to the MAG, the 2 strains were not lacking the same genes. The strain 8825-B1 and MAG 7010000598825.5 shared the genes MAG7010000598825.5_00343, which encoded a protein that was previously annotated as hypothetical protein BIFADO_00229 (100.00% BLASTp identity, accession number EDN83323.1), and MAG7010000598825.5_01204 that encoded a helix-turn-helix domain-containing protein (100.00% BLASTp identity, accession number WP_055680042.1). Both of these genes were absent in strain 8825-Y2b. Whereas the strain 8825-B1 lacked the genes MAG7010000598825.5_01418 and MAG7010000598825.5_00215 that were shared with the strain 8825-Y2b (Fig. 5b). These genes encoded a DUF2207 domain-containing protein (99.59% BLASTp identity, accession number WP_085380746.1) and a protein that was previously annotated as hypothetical protein AD0028_2038 (99.14% BLASTp identity, accession number OSG94723.1). These results show that both strains only have minor functional difference based on comparison of the gene presence. Taking all this together, bacterial strains that closely resemble specific MAGs can be isolated from the same sample.

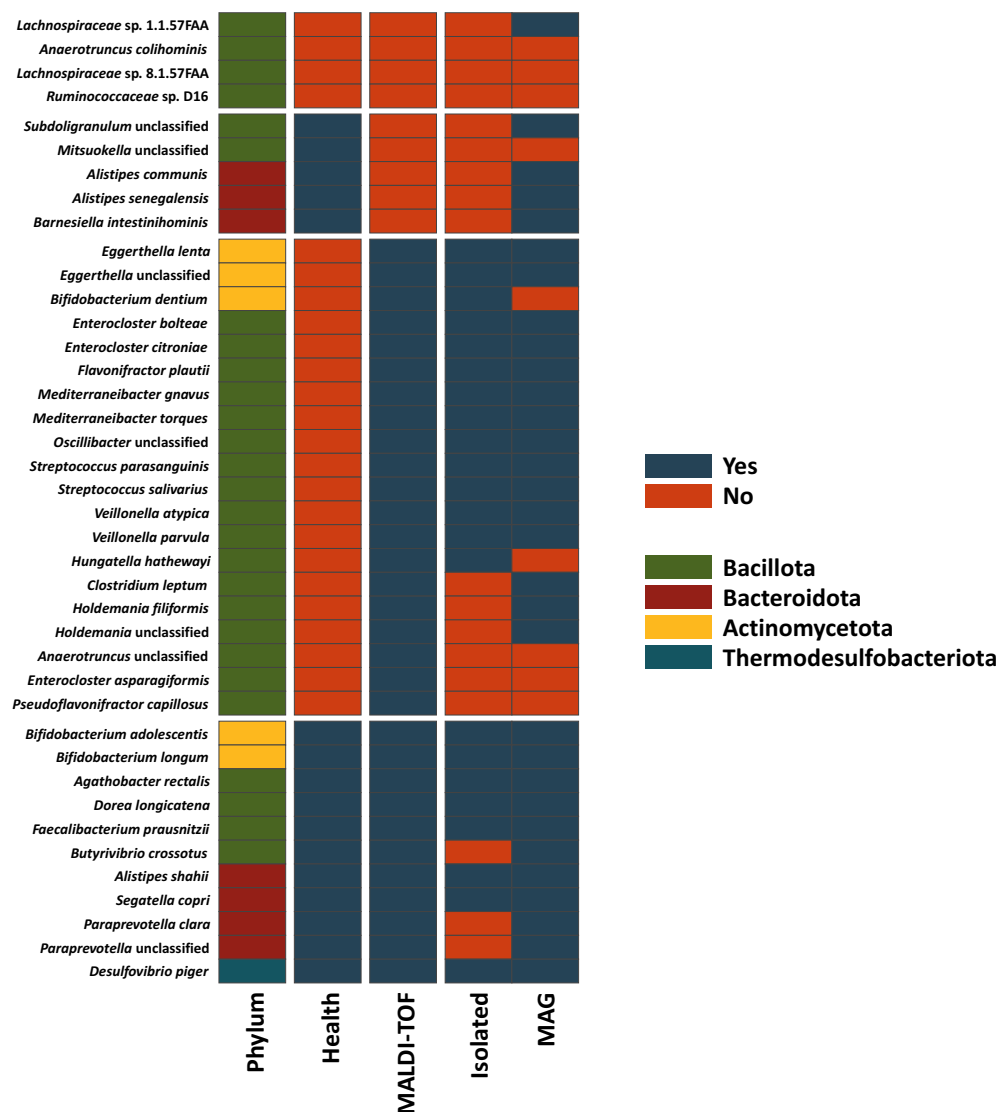
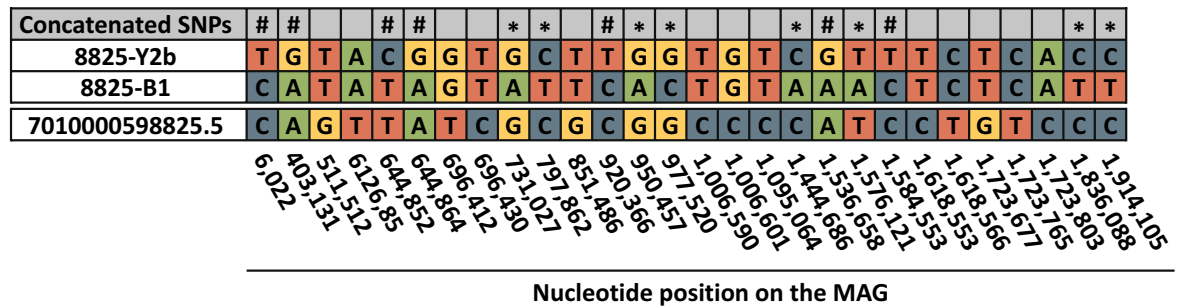


Fig. 4. Culture-dependent and culture-independent identification of top 40 bacterial species that have the strongest associated with the host health status. The phyla of the genera are depicted in green for the *Bacillota*, dark red for the *Bacteroidota*, yellow for the *Actinomycetota* and dark blue for the *Thermodesulfobacteriota*. The association of the bacterial MAGs with the host health (Yes: associated with health, No: associated with disease, as described by Gacesa et al.¹⁵), the presence or absence of reference spectra in the MALDI TOF MS databases, the presence or absence in the culture collection that was isolated in this study, and the identified MAGs are indicated (Yes; blue, No; red). The bacteria are grouped based on the presence or absence in the MALDI-TOF MS databases and the host health status.

Discussion

This study presents a culture collection of 2,985 human gut bacterial isolates from fecal samples of both non-IBD individuals and 1,362 bacterial isolates from fecal samples of IBD patients. These participants are enrolled in large cohort studies, and therefore, a wealth of metadata is available for these subjects, including health status parameters. Combining this data enables more accurate research into the community dynamics of gut bacterial strains within specific species, and the interaction between the host and strains that originate from participants who are non-IBD or suffer from IBD. This is of special interest because improved sequencing technologies and large cohort studies provide promising leads that reveal, with considerable statistical strength, the association of specific bacteria with host health or disease. The sequencing techniques have improved immensely over time, which has allowed much more precise associations of variations in the gut microbiome with health and IBD (e.g., species or even strain level instead of family or genus level). Recently, Kumbhari et al. described different bacterial strains that were associated with IBD and seemed to have adapted to the intestinal environment as created during active disease and were postulated to outcompete other strains or possibly drive or maintain active disease²⁰. Large metagenomic studies, such as the aforementioned study performed by Kumbhari et al. provides an extensive insight into the potential functional role of gut bacteria in the host health status and

a



b

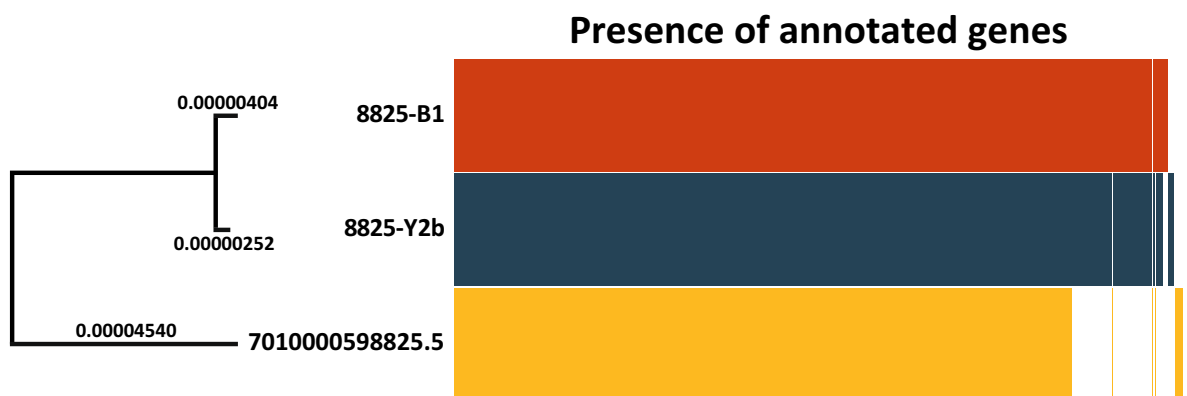


Fig. 5. Genomic comparison of isolates and the associated metagenome-assembled genome **a)** alignment of the concatenated SNPs of the 2 isolates of *B. adolescentis* with the MAG that was obtained from the corresponding participant as reference sequence. The isolate 8825-Y2b differed 20 SNPs and the isolate 8825-B1 differed 21 SNPs to the MAG, based on the covered positions. The nucleotide positions that resemble the reference are displayed with * for 8825-Y2b and with # for 8825-B1. **b)** The phylogenetic distance between the isolates and that MAG is shown on the left. This was calculated with the maximum likelihood method (Tamura-Nei model) based on the core genome. The tree is drawn to scale with the branch length measured in the relative number of substitutions per site, as displayed next to the branches. On the right, the presence of annotated genes is schematically displayed for isolate 8825-B1 (red), 8825-Y2b (blue) and the MAG (yellow). Each annotated gene is depicted as a line, showing 1,548 overlapping genes and 21 genes that were annotated in the MAG, but missing in the isolates.

during various stages of IBD. However, experimental validation with specific bacterial strains still needs to be performed. The need for experimental validation raises the need for isolation of specific bacterial strains that are linked to disease-specific associations within those cohorts. The culture collection in the present study includes a targeted isolation effort to isolate specific strains of *B. adolescentis*, which is a bacterial species that is strongly positively associated with health of the host, as described within the same cohort as used for this culturomics effort¹⁵. The targeted isolation yielded strains that resemble the identified MAG that was obtained from the same participant, showcasing the potential of population and patient biobanks to provide strains that originate from hosts with specific health statuses.

The amount of bacterial isolates originating from non-IBD participants differed from those of IBD patients. The culture media in the current study were mainly based on YCFA as basal medium, which is a medium that is designed for fastidious bacteria belonging to the *Oscillospiraceae* and *Lachnospiraceae*^{27,45}. These bacteria are generally reduced in the gut microbiota of IBD patients, which is supported by previous reports concerning the metagenomic data used in this study¹⁵. The reduced relative abundance of these bacterial species potentially explains the lower number of isolates from IBD patients in this study^{14,15}.

Identification of the isolates in this study was performed with MALDI-TOF MS, providing fast and accurate identification of isolates and enabling a high-throughput approach of culturomics in a relatively simple manner. However, identification relies on reference databases, which are often commercially available databases in which clinically less relevant species are most often underrepresented and therefore are less frequently or not at all identified. Efforts have already been made to expand the catalogue of available reference spectra of such

bacteria, including members of the families *Clostridiaceae*, *Lachnospiraceae* and *Oscillospiraceae*^{27,46,47}, but the expansion of reference databases is an ongoing process, and the lack of reference spectra remains a limitation for the identification of isolates with MALDI-TOF MS. Nevertheless, the reference databases that were used in this study contained already 69 genera out of the 131 that were identified based on MAGs within the cohort. The culture collection, which is presented in this study, includes 61 of the genera that were present in the MALDI-TOF database. The culturomics effort also resulted in the isolation of 23 genera that were not detected based on MAGs. This was especially the case for genera belonging to *Actinomycetota*, which could be due to the sensitive media we used for the oxygen-tolerant bacteria (e.g. *Micrococcaceae*). The abundance of these genera could be below the threshold for metagenomic read filtering or a potential bias due to different DNA extraction techniques, resulting in absence of these genera in the metagenomic data. In contrast, 9 genera were detected based on MAGs, but were not isolated on the culture media in this study. A previous culturomics effort managed to isolate 4 out of these 9 genera using YCFAG or YCFAP medium²⁷, showing that these 4 genera are culturable on the media in the present study. The other 5 genera; *Bilophila*, *Dialister*, *Coprobacter*, *Paraprevotella*, and *Parasutterella*, should be cultured on more specific media. *Bilophila* and *Dialister* can be cultured on BBA, PYG medium or FAA medium supplemented with blood, as described in the original isolations^{48,49}. These media resemble the FSM or YCFAG that were used in this study, but specific growth tests with these species or analysis of their genome, as recently proposed⁵⁰, should reveal the suitability of the media. However, we expect growth of these species on media that were used in this study, and although the FSM was supplemented with josamycin, vancomycin, and neomycin, it is recorded that these antibiotics are not effective against these anaerobic gram-negative bacilli^{51,52}. *Coprobacter* is a genus of gram-negative bacteria that can be cultured on Eggerth-Gagnon medium supplemented with hemin^{53,54}. YCFAG medium closely resembles this medium, but lacks soluble starch and products that resemble meat extract, such as peptone or Lab Lemco powder, although does YCFAG medium contain casitone. *Paraprevotella* and *Parasutterella* both grow on modified Gifu anaerobic medium, which also resembles YCFAG, but also differs in the presence of meat extract^{55–57}. Most media that were used in this study were non-selective and therefore allowed the growth of a high diversity of bacteria, which could have resulted in loss of bacteria that had a low relative abundance in the sample or were outcompeted. Indeed, the relative abundance of *Bilophila*, *Dialister*, *Coprobacter*, *Paraprevotella*, and *Parasutterella* was between 0.004% and 0.078% based on the metagenomic data of these fecal samples, which was notably lower than the average relative abundance of a bacterial species in these samples, which was 1.521%.

The discrepancy between the identification results of the isolates and the MAGs within the same cohorts highlights that ideally multiple culture-dependent and culture-independent approaches are implemented, because these approaches complement each other in order to increase the identification rate. A major drawback of non-automated culturomics is the workload that increases non-linear when multiple culture conditions are included. A more practical approach instead of culturing all samples is selecting specific samples for culturing based on sequencing results, thereby isolating low abundant species and multiple strains. This will enhance the detection of the actual bacterial richness within samples and thereby helps in a better understanding of the individual contribution of bacteria to health or IBD. Furthermore, the genome sequences of the isolates within this culture collection can act as scaffolds for genome assembly and therefore will benefit future MAG assembly. This will result in more accurate MAGs and that would allow a better prediction for subsequent sample selection based on MAGs, having a synergetic effect. The culture collection that is presented in this study contains bacterial isolates that are strongly associated with health or IBD, as also described in the same cohort¹⁵. These isolates thus open a promising avenue for the validation of associations that are identified with the use of metagenomic data.

Data availability

The raw microbiome sequencing data, processed microbiome data and basic host phenotypes of the Lifelines DMP are available at the European Genome-Phenome Archive under the study accession EGAS00001005027. The raw microbiome sequencing data of the RISE-UP participants that are used in this study is available at the European Genome-Phenome Archive under the study accession EGAS50000000982. The commercial MALDI-TOF MS database used in this study is available at Bruker Daltonics, which was used under license for the current study, and is not publicly available. The MaldiGut database is published previously (<https://doi.org/10.1163/18762891-bja00042>) and is publicly available at <https://github.com/NPIMP/MaldiGut>.

Code availability

The underlying code (and training/validation datasets) for this study is available in GitHub and can be accessed via this link https://github.com/GRONINGEN-MICROBIOME-CENTRE/GMH_MGS_pipeline

Received: 11 June 2025; Accepted: 14 November 2025

Published online: 27 November 2025

References

1. Gomaa, E. Z. Human gut microbiota/microbiome in health and diseases: A review. *Antonie Van Leeuwenhoek* **113**, 2019–2040 (2020).
2. Round, J. L. & Mazmanian, S. K. The gut microbiota shapes intestinal immune responses during health and disease. *Nat. Rev. Immunol.* **9**, 313–323 (2009).
3. Tlaskalová-Hogenová, H. et al. The role of gut microbiota (commensal bacteria) and the mucosal barrier in the pathogenesis of inflammatory and autoimmune diseases and cancer: Contribution of germ-free and gnotobiotic animal models of human diseases. *Cell. Mol. Immunol.* **8**, 110–120 (2011).
4. Dethlefsen, L., Huse, S., Sogin, M. L. & Relman, D. A. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol.* **6**, e280 (2008).

5. Leviatan, S., Shoer, S., Rothschild, D., Gorodetski, M. & Segal, E. An expanded reference map of the human gut microbiome reveals hundreds of previously unknown species. *Nat. Commun.* **13**, 3863 (2022).
6. Qin, J. et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
7. Tap, J. et al. Towards the human intestinal microbiota phylogenetic core. *Environ. Microbiol.* **11**, 2574–2584 (2009).
8. Wen, M. et al. Diversity and host interaction of the gut microbiota in specific pathogen-free pigs. *Front. Microbiol.* **15**, 1402807 (2024).
9. Chow, J., Lee, S. M., Shen, Y., Khosravi, A. & Mazmanian, S. K. Host-Bacterial symbiosis in health and disease. *Adv. Immunol.* **107**, 243–274 (2010).
10. Tsavkelova, E. A., Botvinko, I. V., Kudrin, V. S. & Oleskin, A. V. Detection of neurotransmitter amines in microorganisms with the use of high-performance liquid chromatography. *Dokl. Biochem.* **372**, 115–117 (2000).
11. Kamada, N., Seo, S.-U., Chen, G. Y. & Núñez, G. Role of the gut microbiota in immunity and inflammatory disease. *Nat. Rev. Immunol.* **13**, 321–335 (2013).
12. Panwar, R. B., Sequeira, R. P. & Clarke, T. B. Microbiota-mediated protection against antibiotic-resistant pathogens. *Genes Immun.* **22**, 255–267 (2021).
13. Alatab, S. et al. The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol. Hepatol.* **5**, 17–30 (2020).
14. Alam, M. T. et al. Microbial imbalance in inflammatory bowel disease patients at different taxonomic levels. *Gut Pathogens* **12**, 1 (2020).
15. Gacesa, R. et al. Environmental factors shaping the gut microbiome in a Dutch population. *Nature* **604**, 732–739 (2022).
16. Vich Vila, A. et al. Gut microbiota composition and functional changes in inflammatory bowel disease and irritable bowel syndrome. *Sci. Transl. Med.* **10**, eaap8914 (2018).
17. Lo Sasso, G. et al. Inflammatory bowel disease-associated changes in the gut: Focus on kazan patients. *Inflamm. Bowel Dis.* **27**, 418–433 (2020).
18. Franzosa, E. A. et al. Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat. Microbiol.* **4**, 293–305 (2019).
19. Sommersberger, S. et al. Altered fecal bile acid composition in active ulcerative colitis. *Lipids Health Dis.* **22**, 199 (2023).
20. Kumbhari, A. et al. Discovery of disease-adapted bacterial lineages in inflammatory bowel diseases. *Cell Host Microbe* **32**, 1147–1162.e12 (2024).
21. Browne, H. P. et al. Culturing of ‘unculturable’ human microbiota reveals novel taxa and extensive sporulation. *Nature* **533**, 543–546 (2016).
22. Forster, S. C. et al. A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat. Biotechnol.* **37**, 186–192 (2019).
23. Lagier, J.-C. et al. Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat. Microbiol.* **1**, 16203 (2016).
24. Li, C. et al. Deep insights into the gut microbial community of extreme longevity in south Chinese centenarians by ultra-deep metagenomics and large-scale culturomics. *NPJ Biofilms Microbiomes* **8**, 28 (2022).
25. Mukhopadhyay, I. et al. Comparison of microbial signatures between paired faecal and rectal biopsy samples from healthy volunteers using next-generation sequencing and culturomics. *Microbiome* **10**, 171 (2022).
26. Poyet, M. et al. A library of human gut bacterial isolates paired with longitudinal multiomics data enables mechanistic microbiome research. *Nat. Med.* **25**, 1442–1452 (2019).
27. Plomp, N. et al. A convenient and versatile culturomics platform to expand the human gut culturome of Lachnospiraceae and Oscillospiraceae. *Benef. Microbes* <https://doi.org/10.1163/18762891-bja00042> (2024).
28. von Martels, J. Z. H. et al. Riboflavin supplementation in patients with crohn’s disease [the RISE-UP study]. *J. Crohns Colitis* **14**, 595–607 (2020).
29. Maaser, C. et al. ECCO-ESGAR Guideline for Diagnostic Assessment in IBD Part 1: Initial diagnosis, monitoring of known IBD, detection of complications. *J. Crohns Colitis* **13**, 144–164K (2019).
30. Tigchelaar, E. F. et al. Cohort profile: LifeLines DEEP, a prospective, general population cohort study in the northern Netherlands: Study design and baseline characteristics. *BMJ Open* **5**, e006772 (2015).
31. Veloo, A. C. M. et al. The optimization and validation of the Biotyper MALDI-TOF MS database for the identification of Gram-positive anaerobic cocci. *Clin. Microbiol. Infect.* **22**, 793–798 (2016).
32. Schuurman, T., de Boer, R. F., Kooistra-Smid, A. M. D. & van Zwet, A. A. Prospective study of use of PCR amplification and sequencing of 16S ribosomal DNA from cerebrospinal fluid for diagnosis of bacterial meningitis in a clinical setting. *J. Clin. Microbiol.* **42**, 734–740 (2004).
33. Cole, J. R. et al. Ribosomal Database Project: Data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **42**, D633–D642 (2014).
34. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naïve bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
35. Yoon, S.-H. et al. Introducing EzBioCloud: A taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* **67**, 1613–1617 (2017).
36. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
37. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
38. Truong, D. T. et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat. Methods* **12**, 902–903 (2015).
39. Yoon, S.-H., Ha, S.-M., Lim, J., Kwon, S. & Chun, J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* **110**, 1281–1286 (2017).
40. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
41. Page, A. J. et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691–3693 (2015).
42. Felsenstein, J. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J. Mol. Evol.* **17**, 368–376 (1981).
43. Tamura, K. & Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**, 512–526 (1993).
44. Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).
45. Duncan, S. H., Hold, G. L., Harmsen, H. J. M., Stewart, C. S. & Flint, H. J. Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. *Int. J. Syst. Evol. Microbiol.* **52**, 2141–2146 (2002).
46. Asare, P. T. et al. A MALDI-TOF MS library for rapid identification of human commensal gut bacteria from the class Clostridia. *Front. Microbiol.* **14**, 1104707 (2023).
47. URMS Data Base – IHU. <https://www.mediterranean-infection.com/acces-ressources/base-de-donnees/urms-data-base/> (2023).
48. Sakamoto, M. et al. *Dialister hominis* sp. nov., isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* **70**, 589–595 (2020).
49. Downes, J., Munson, M. & Wade, W. G. *Dialister invisus* sp. nov., isolated from the human oral cavity. *Int. J. Syst. Evol. Microbiol.* **53**, 1937–1940 (2003).

50. Armetta, J., Li, S. S., Vaaben, T. H., Vazquez-Urbe, R. & Sommer, M. O. A. Metagenome-guided culturomics for the targeted enrichment of gut microbes. *Nat. Commun.* **16**, 663 (2025).
51. Hu, Y., Liu, L., Zhang, X., Feng, Y. & Zong, Z. In Vitro Activity of Neomycin, Streptomycin, Paromomycin and Apramycin against Carbapenem-Resistant Enterobacteriaceae Clinical Strains. *Front. Microbiol.* **8**, 2275 (2017).
52. Vardanyan, R. S. & Hruby, V. J. 32 - Antibiotics. in *Synthesis of Essential Drugs* (eds. Vardanyan, R. S. & Hruby, V. J.) 425–498 (Elsevier, Amsterdam, 2006). <https://doi.org/10.1016/B978-044452166-8/50032-7>.
53. Sakamoto, M. et al. *Coprobacter secundus* subsp. *similis* subsp. nov. and *Solibaculum mannosilyticum* gen. nov., sp. nov., isolated from human feces. *Microbiol. Immunol.* **65**, 245–256 (2021).
54. Shkoporov, A. N. et al. *Coprobacter fastidiosus* gen. nov., sp. nov., a novel member of the family Porphyromonadaceae isolated from infant faeces. *Int. J. Syst. Evol. Microbiol.* **63**, 4181–4188 (2013).
55. Morotomi, M., Nagai, F. & Watanabe, Y. *Parasutterella secunda* sp. nov., isolated from human faeces and proposal of Sutterellaceae fam. nov. in the order Burkholderiales. *Int. J. Syst. Evol. Microbiol.* **61**, 637–643 (2011).
56. Morotomi, M., Nagai, F., Sakon, H. & Tanaka, R. *Paraprevotella clara* gen. nov., sp. nov. and *Paraprevotella xylaniphila* sp. nov., members of the family 'Prevotellaceae' isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* **59**, 1895–1900 (2009).
57. Nagai, F., Morotomi, M., Sakon, H. & Tanaka, R. *Parasutterella excrementihominis* gen. nov., sp. nov., a member of the family Alcaligenaceae isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* **59**, 1793–1797 (2009).
58. Kamachi, K., Koide, K., Otsuka, N., Goto, M. & Kenri, T. Whole-Genome Analysis of *Bordetella pertussis* MT27 Isolates from School-Associated Outbreaks: Single-Nucleotide Polymorphism Diversity and Threshold of the Outbreak Strains. *Microbiol. Spectr.* **11**, e0406522 (2023).

Acknowledgements

This study was funded by the collaborative TIMID project (LSHM18057-SGF) financed by the PPP allowance made available by Top Sector Life Sciences & Health to the Samenwerkende Gezondheidsfondsen (SGF) to stimulate public–private partnerships and co-financing by health foundations that are part of the SGF. NP was supported by a grant from the Graduate School of Medical Sciences of the University of Groningen, the Netherlands. RKW received speaker fee from AbbVie, MSD, Olympus, AstraZeneca and Ferring, is supported by the Seerave Foundation, the Netherlands Organization for Scientific Research (NWO; 016.136.308), the Dutch Digestive Foundation under Grant MLDS D16–14 and the EU Horizon Europe Program grant miGut-Health: personalized blueprint of intestinal health (101095470) and received grants from Takeda Pharmaceuticals and Johnson and Johnson. SW, CW and JS are supported by The Netherlands Organ-on-Chip Initiative, a Netherlands Organization for Scientific Research (NWO Gravitation project 024.003.001), funded by the Ministry of Education, Culture, and Science of the Government of the Netherlands. The funders played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript. The authors thank A. Brouwer, M. de Vries and S. Meis for their technical assistance.

Author contributions

NP designed and performed the culturomics pipeline, data collection, formal analysis, data visualization and writing of the original draft of the manuscript. RG and JS analyzed the metagenomic data. HJMH, JNS, CW, SW and RKW conceived and coordinated the study. All authors critically revised and approved the manuscript.

Funding

Samenwerkende Gezondheidsfondsen, LSHM18057-SGF, Graduate School of Medical Sciences of the University of Groningen, the Netherlands, Nederlandse Organisatie voor Wetenschappelijk Onderzoek: NWO Gravitation project 024.003.001, 016.136.308, Abbvie, MSD, Olympus, AstraZeneca, Ferring Research Institute, Seerave Foundation, Maag Lever Darm Stichting MLDS D16–14, HORIZON EUROPE Framework Programme 101095470, Takeda Pharmaceuticals, Johnson and Johnson.

Declarations

Competing interests

The authors declare no competing interests.

Consent to participate

Informed consent was obtained from all subjects and/or their legal guardian(s).

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-29138-4>.

Correspondence and requests for materials should be addressed to H.J.M.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025