environmental microbiology

Environmental Microbiology (2022) **24**(9), 3861–3881 doi:10.1111/1462-2920.15935

Anaerobic single-cell dispensing facilitates the cultivation of human gut bacteria

Afrizal Afrizal[1], **Thomas C. A. Hitch**[1], **Alina Viehof**[1], **Nicole Treichel**[1], **Thomas Riedel**[2,3], **Birte Abt**[2,3], **Eva M. Buhl**[4], **Dietrich Kohlheyer**[5,6], **Jörg Overmann**[2,3,7] and **Thomas Clavel**[1*]

[1]*Functional Microbiome Research Group, Institute of Medical Microbiology, University Hospital of RWTH Aachen, Aachen, Germany.*

[2]*Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany.*

[3]*German Center for Infection Research (DZIF), Partner site Hannover-Braunschweig, Braunschweig, Germany.*

[4]*Electron Microscopy Facility, Institute of Pathology, University Hospital of RWTH Aachen, Aachen, Germany.*

[5]*Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich, Jülich, Germany.*

[6]*Aachener Verfahrenstechnik (AVT-MSB), RWTH Aachen University, Aachen, Germany.*

[7]*Institute of Microbiology, Technical University of Braunschweig, Braunschweig, Germany.*

Summary

Cultivation via classical agar plate (CAP) approaches is widely used to study microbial communities, but they are time-consuming. An alternative approach is the application of single-cell dispensing (SCD), which allows high-throughput, label-free sorting of microscopic particles. We aimed to develop a new anaerobic SCD workflow to cultivate human gut bacteria and compared it with CAP using faecal communities on three rich culture media. We found that the SCD approach significantly decreased the experimental time to obtain pure cultures from 17 ± 4 to 5 ± 0 days, while the isolate diversity and relative abundance coverage were comparable for both approaches. We further tested the total captured fraction by sequencing the sorted bacteria directly after growth as bulk biomass from 2400 dispensed single cells without downstream identification of individual strains. In this approach, the cultured fraction increased from 35.2% to 52.2% for SCD, highlighting the potential for deeper cultivation projects from single samples. SCD-based cultivation also captured species not detected by sequencing (16 ± 5 per sample, including seven novel taxa). From this work, 82 human gut bacterial species across five phyla (Actinobacteriota, Bacteroidota, Desulfobacterota, Firmicutes and Proteobacteria) and 24 families were obtained, including the first cultured member of 11 novel genera and 10 novel species that were fully characterized taxonomically.

Received 24 October, 2021; revised 3 February, 2022; accepted 8 February, 2022. *For correspondence. E-mail tclavel@ukaachen.de; Tel. +49 241 80 855 23.

Introduction

Microbiota are complex communities of microbes (Berg *et al*., 2020) that are important for nutrient cycling in the environment (Bahram *et al*., 2018; Salazar *et al*., 2019) or for the host species they colonize (Lynch and Hsiao, 2019). The gut microbiome has been extensively studied over the last 20 years thanks to high-throughput sequencing technologies (Eckburg *et al*., 2005; Knight *et al*., 2018) and is known to modulate multiple host functions such as metabolic and immune responses and thereby influence the development and treatment of infections and chronic diseases (Rutgeerts *et al*., 1991; van Nood *et al*., 2013; Sonnenburg *et al*., 2016; Janney *et al*., 2020). The most dominant members of the human gut microbiota are bacteria, with an estimated number of 4×10^{13} cells per individual representing a few hundred grams of biomass and a diversity of 300–400 species per human individual (Moore and Holdeman, 1974; Stephen and Cummings, 1980; Clavel *et al*., 2016; Sender *et al*., 2016). Unfortunately, a substantial fraction of these gut bacteria remains unknown. This significantly hampers our understanding of microbe–host interactions and the implementation of microbiota-based therapeutics. The unknowns within gut microbiota include bacterial taxa that have so far evaded detection (i.e. not yet captured by any method) or those that have been captured by sequencing but not yet cultured (Thomas and Segata, 2019). Even though estimates of the cultured fraction of gut microbiota vary greatly depending on the

host species considered and the methodology used, 30%–65% of the dominant bacterial diversity detected by sequencing in the mouse or human gut is covered by cultured taxa (Lagkouvardos *et al*., 2017b; Almeida *et al*., 2019; Almeida *et al*., 2021; Hitch *et al*., 2021a). Hence, whilst a considerable fraction of gut microbial communities is already cultured, more effort must be put into isolating and characterizing gut microbes to allow further progress in microbiome research.

The research community is slowly catching up with the cultivation of bacteria from the intestine of several mammalian species (Browne *et al*., 2016; Lagier *et al*., 2016; Lagkouvardos *et al*., 2016b; Seshadri *et al*., 2018; Zehavi *et al*., 2018; Forster *et al*., 2019; Zou *et al*., 2019; Wylensek *et al*., 2020; Beresford-Jones *et al*., 2022) and state-of-the-art projects have been launched to preserve and explore bacterial diversity (Bello *et al*., 2018; Groussin *et al*., 2021). However, while many papers report on the size and novelty of their culture collections, few make the effort to fully describe and validate their novel taxa. This is mostly due to underfunded international culture collections and correspondingly the tedious process of depositing novel isolates and validating their names (Murray *et al*., 2020). Moreover, contemporary cultivation approaches are very similar to the work done after the advent of anaerobic technologies in the 1960s (Floch *et al*., 1968; Attebery and Finegold, 1969; Overmann, 2013). The higher output of projects nowadays is mostly a consequence of increasing manpower to parallelise isolation using a range of agar media combined with more rapid and accurate identification of isolates by mass spectrometry and whole-genome sequencing. In contrast, innovative isolation procedures with higher throughput are rare (Overmann *et al*., 2017). In 2011, a pioneering study by Goodman *et al*. (2011) proposed a dilution-to-extinction approach to create personal culture collections. However, mixed cultures may be detected late in such a workflow and few follow-up studies have been performed (Faith *et al*., 2013), where the isolation of <50 species per sample from a limited number of samples was reported. Efficient separation of cells from complex microbial communities can be achieved by droplet microfluidics. Such approach has been used successfully for the isolation of bacteria from the environment (Hu *et al*., 2020) and under anaerobic conditions from the intestine of termites (Zhou *et al*., 2019) and humans (Watterson *et al*., 2020). In the latter study, the cultivation system used is not commercially available, preventing a broader application of the system. Furthermore, it is unable to provide pure isolates for deposit in culture collections.

The aim of the present work was to benchmark a new cultivation approach based on single-cell dispensing (SCD) (Riba *et al*., 2016) for isolation of strictly anaerobic bacteria from human stool. The speed and efficacy of the method were compared to a classical agar plate (CAP) workflow. The respective cultured fractions were explored via 16S rRNA gene amplicon sequencing and new taxonomic diversity was recovered and described.

Results

General features of the single-cell dispensing workflow

We first assessed the speed and overall output of the SCD workflow based on the analysis of pure isolates compared with CAP-based isolation (Fig. 1A).

The SCD approach required slightly more hands-on time to process the samples: 5.0 $\pm$ 0.3 h per sample until the 2400 single cells were dispensed (with addition of the 480 control wells) versus 3.7 $\pm$ 0.4 h until all dilutions were plated and colonies then re-streaked in the CAP approach ($n = 4$; $p = 0.0286$; Mann–Whitney U test). However, this processing time represented only a minor fraction of the 414 $\pm$ 99.5 h (13–21 days) required for incubation for the CAP method versus 120 h only (5 days) until pure cultures were obtained in the SCD method (Fig. 1B). Moreover, depth of analysis (i.e. number of dispensed single cells) can be easily scaled up without a substantial rise in hands-on time. Increasing the number of dispensed cells is also necessary to capture low abundant bacteria, as the stochastic nature of cells placed within droplets means low abundant members of the microbiota are dispensed less often.

Next, we determined the proportion of droplets leading to colony formation, hereon referred to as growth output. For the entire series of experiments (Fig. 1A), 120 agar plates (30 plates per sample; 10 for each culture medium) were used. The average growth output per condition (i.e. sample/medium pair; 800 single cells) was 33 $\pm$ 11% ($n = 4$ samples) (Fig. 1C). Values ranged from 5% to 61% per plate (80 single-cell droplets) depending on the sample and culture medium. Overall, inter-sample variations in the cultured output were higher than differences between culture media for any given sample (Fig. 1C).

A total of 45 single colonies per cultivation approach (SCD vs. CAP) and samples were further processed, resulting in 360 pure isolates that were subjected to species identification. Only 27% of these isolates, on average, could be directly assigned to taxa included in the MALDI database (Fig. 1D). Hence, the majority required further identification by 16S rRNA gene sequencing. This further step added another 33%–60% of isolates identified as known taxa, whilst the fraction of potentially novel taxa (<98.7% sequence identity to any bacteria with a valid name) ranged from 9% to 31% depending on the

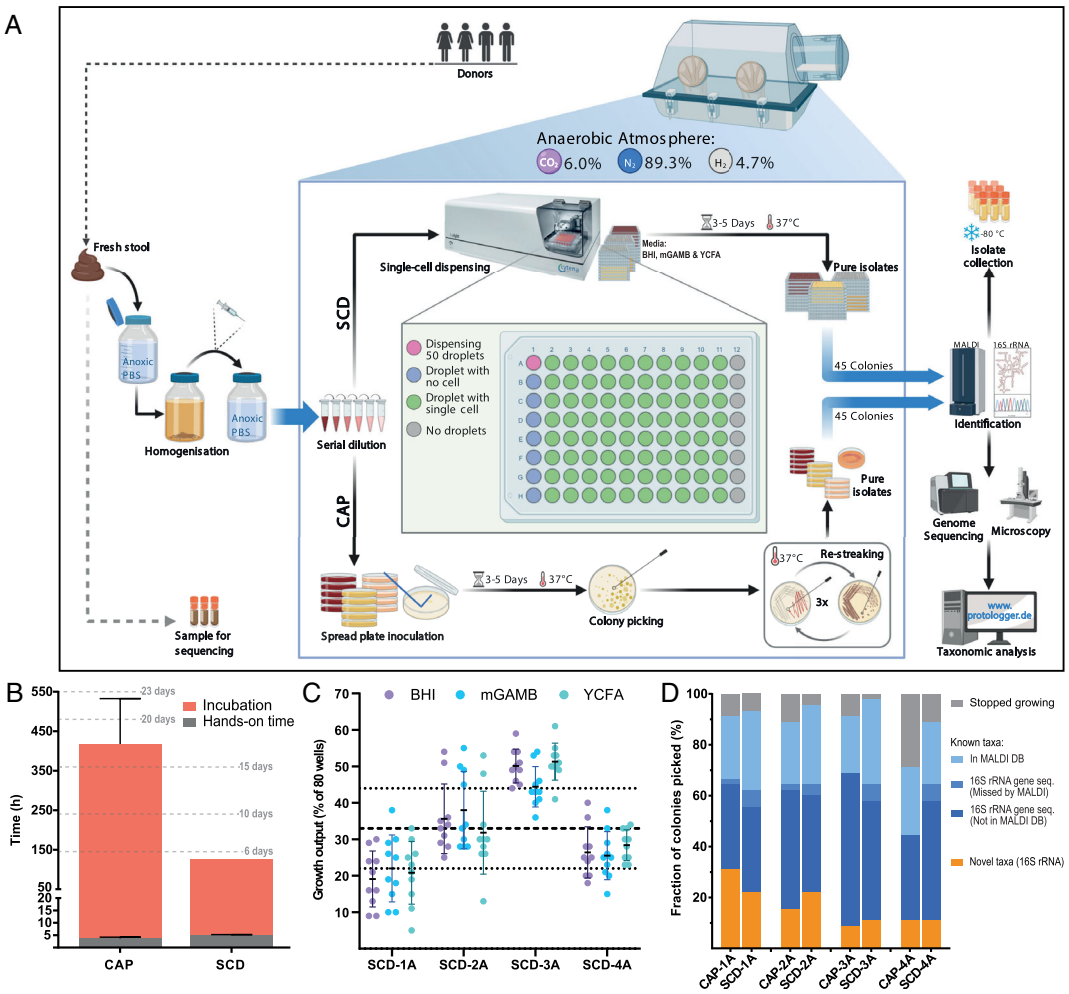

Fig. 1. Comparison between SCD and CAP anaerobic cultivation.
A. Experimental workflow to compare the speed and efficacy of SCD against CAP isolation of pure cultures. Fresh faecal samples were obtained from four different donors. Each sample was processed aseptically following a two-step initial sample preparation to minimise oxygen exposure outside the anaerobic workstation (left-hand side of the scheme). The diluted samples were subsequently transferred into the anaerobic workstation accommodating the single-cell dispenser and then used as shared starting material for both the SCD and CAP arm of the studies. This figure panel was created in BioRender.
B. Time required by each approach, from sample preparation to obtaining pure isolates.
C. Proportion of dispensed single-celled droplets grown to colony, referred to as growth output. Each dot represents one plate (80 single-cell droplets), with a total of 30 plates (10 per medium) for each sample (SCD-1A to -4A). The dashed line is the mean growth output for 800 droplets (10 plates), the dotted lines representing standard deviation.
D. Distribution of picked colonies ($n = 45$ per sample per cultivation approach) based on the identification method and identity categories. The category 'Missed by MALDI' means that isolates were identified by 16S rRNA gene sequencing as species included in the MALDI database (DB), but the first MALDI measurement failed to deliver a confident identification. The bacteria that could not be maintained in culture during re-streaking (CAP approach) or in Hungate tubes (CAP and SCD approach) were referred to as 'Stopped growing'.

sample and culture medium. The number of novel taxa obtained via the SCD or CAP approach was identical, with an average of four new species per condition. Of note, due to the multiple re-streaking steps and incubation periods required prior to final storage of the isolates in the CAP approach, a higher fraction of picked colonies stopped growing compared with SCD, although this difference did not reach statistical significance (14.5 ± 8.4% vs. 6.1 ± 3.3%; $p = 0.1429$; Mann–Whitney *U* test).

Isolate-based cultured fractions

In order to determine the fraction of bacteria isolated from the cultivation experiment, we performed 16S rRNA gene amplicon sequencing on the original faecal samples. A total of 125 molecular species (71 ± 8 OTUs per sample), occurring at a relative abundance ≥0.1% in at least one sample were detected across all four faecal samples analysed, with an average of 34 954 ± 11 274 high-quality, assembled sequences per sample (Supplementary

Table S1 available in the project-specific data repository: https://github.com/ClavelLab/SCD). More than half (53%) of the OTUs detected within a sample could be obtained by cultivation, 35.2% via the SCD approach alone at the selected depth of 45 colonies per sample (180 isolates in total) (Fig. 2A). The diversity of isolates obtained per sample was 21 ± 2 species on average and similar between both methods. The contribution of each of the SCD or CAP approach to the cultured fraction was comparable, with 18 and 23 uniquely isolated species respectively. The species missed by cultivation were predominantly taxa with a low relative abundance (<1.0%) within the phyla Firmicutes and Proteobacteria. The isolate-based cultured fraction represented a total relative abundance of 62.8 ± 19.0% sequence-based diversity. On a sample-by-sample basis, the majority of abundant OTUs (>1% relative abundance) was captured by cultivation, with the exception of faecal sample 1A (Fig. 2B). Whilst a substantial number of low abundant OTUs remained not yet cultured (grey bars in Fig. 2B, top panel), seven isolates were not detected by sequencing, including four novel taxa (golden star with strain designations in Fig. 2A). Of the 30 OTUs shared across all four faecal samples, designated 'core OTUs', 87% were captured by cultivation (Fig. 2A).

From all four human faecal samples, a total of 82 bacterial species could be isolated during this cultivation work (including both methods). Information about this collection as well as sequences and accession numbers are available in Supplementary Table S2 within the project-specific data repository (https://github.com/ClavelLab/SCD). The isolates spanned five phyla, mostly Firmicutes (66%) and Bacteroidota (21%), and 24 families, dominated by *Lachnospiraceae* (37%), *Bacteroidaceae* (13%) and *Oscillospiraceae* (valid name; *Ruminococcaceae* as heterotypic synonym in SILVA) (9%). The collection includes cultured members of 11 genera and 10 novel species without a valid name at the time of re-submission, for which draft genomes were generated. These novel taxa, which include dominant members within the families *Lachnospiraceae* and *Oscillospiraceae* and six high-priority, most wanted taxa according to Human Microbiome Project definition (Fodor *et al*., 2012), are all in the process of being deposited at the Leibniz Institute DSMZ for public availability. From these 21 isolates representing novel taxa, 16 (only) are taxonomically described in this manuscript (see protologues at the end of the methods section), as the remaining five have names that are still not-yet validly published but currently in the process of validation. Scanning and transmission electron micrographs for these strains are also available via the aforementioned project data repository.

Bulk cultured fractions

To appreciate the potential of the SCD approach independent of the work-intensive process of identifying single isolates, a second set of experiments was conducted with fresh stool samples from the same four donors (SCD1-4B) to determine the cultured fractions based on bulk amplicon sequencing of all dispensed cells that gave rise to single colonies after 5 days of growth on the plates. A total of 2400 single cells were dispensed per sample onto 10 plates of each the same three culture media as in the first experiments (BHI, mGAMB, YCFA) (30 plates in total per faecal donor). Colonies grown on any given single plate were pooled and processed for high-throughput amplicon sequencing together with the original faecal samples, resulting in a total of 124 samples sequenced. The experimental design is depicted in Fig. 3. The processed sequencing data representing 31 696 ± 10 750 high-quality sequences per sample and a total diversity of 154 OTUs are provided in Supplementary Table S3 within the project data repository (https://github.com/ClavelLab/SCD).

Compared with the SCD global cultured fraction (i.e. the species isolated from all four samples by SCD approach) in experiment type 1 (SCD vs. CAP comparison based on isolates) that yielded 35.2% of the total faecal OTU richness (125 OTUs), bulk sequencing (this experiment type 2) of the cultured biomass directly after single-cell sorting and incubation resulted in a cultured fraction of 52.2% of the total OTU richness (representing overall 82% cumulative relative abundance of the faecal OTUs), with a mean value of 43.5 ± 15.4% per sample (73.5 ± 13.8% relative abundance) (Fig. 4A). Of note, the cultured fraction corresponding to sample 1 increased markedly in this second set of experiments. This is partly due to the fact the two most dominant taxa in this sample (SCD-1A in Fig. 2B; OTU-4 and -5 in Supplementary Table S1) were missed by cultivation in experiment type 1, either because growth conditions were not ideal for those bacteria to grow as pure cultures or because of a too low number of colonies picked. In contrast, these same two species were contained in the bulk cultured biomass in experiment type 2 (OTU-10 and -16 in Supplementary Table S3).

Any additional plate dispensed added new species to the global pool of cultured diversity, resulting in an average sample-specific diversity of 30 ± 11 species per 800 cells for any culture media (range: 15–46 species) (Fig. 4B). One-third of the cultured species were captured by all three media, while each medium contributed 2%–23% unique species depending on the sample (Fig. 4C).

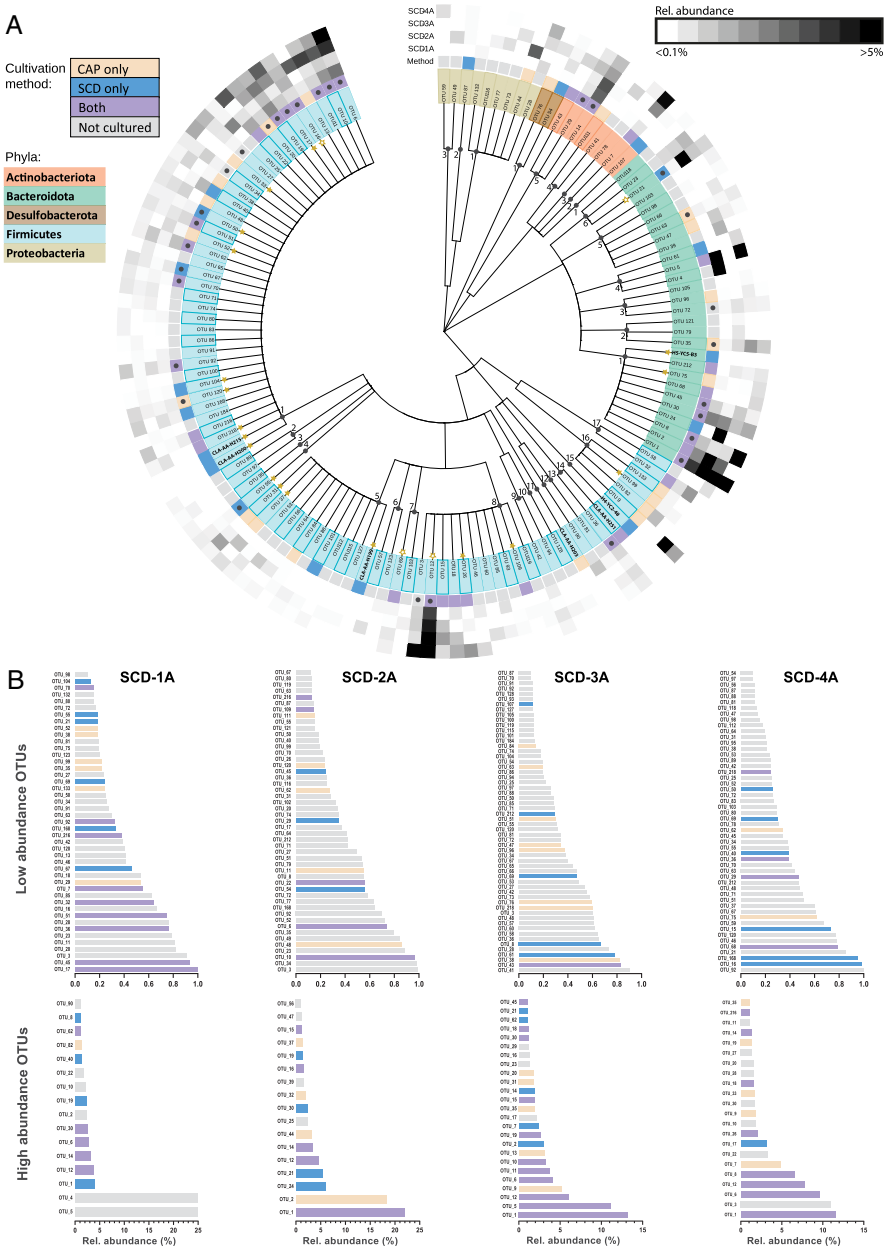

Fig. 2. Comparative analysis between faecal molecular species and cultured isolates.

A. Taxonomic dendrogram of all operational taxonomic units (OTUs) detected in the original stool samples by high-throughput amplicon sequencing. Isolates not detected by amplicon sequencing ($n = 7$) were added to the dendrogram with their respective strain designation written in bold letters. Filled golden stars indicate OTUs that corresponded unambiguously to novel taxa. Empty stars indicate OTUs that shared >97% 16S rRNA gene sequence identity to both novel and known taxa. OTUs are colour-coded by phyla. They were further labelled according to the method used to isolate corresponding cultured species (first outer ring: beige, CAP only; blue, SCD only; violet, obtained by both methods; light grey, not cultured). The black dots in the first outer ring showing the cultivation methods represent the 30 core OTUs, i.e. those present in all four faecal samples. The occurrence of each OTU in each of the four samples analysed is displayed by the grey gradient of relative abundances in the other outer rings. Black dots on the dendrogram represent bacterial families (according to SILVA taxonomy, release 128) designated by an incremental number within each phylum (Actinobacteria: 1. *Atopobiaceae*, 2. *Coriobacteriaceae*, 3. *Coriobacteriales incertae sedis*, 4. *Eggerthellaceae*, 5. *Bifidobacteriaceae*; Bacteroidota: 1. *Bacteroidaceae*, 2. *Barnesiellaceae*, 3. *Marinifilaceae*, 4. *Prevotellaceae*, 5. *Rikenellaceae*, 6. *Tannerellaceae*; Desulfobacterota: 1. *Desulfovibrionaceae*; Firmicutes: 1. *Lachnospiraceae*, 2. *Christensenellaceae*, 3. Unknown *Clostridia* UCG-014; 4. 'Monoglobaceae', 5. *Oscillospiraceae*, 6. *Eubacterium coprostanoligenes* group, 7. *Butyricicoccaceae*, 8. *Ruminococcaceae*, 9. *Peptostreptococcaceae*, 10. *Anaerovoracaceae*, 11. *Peptoniphilaceae*, 12. *Lactobacillaceae*, 13. *Streptococcaceae*, 14. 'Erysipelatoclostridiaceae', 15. *Paenibacillaceae*, 16. *Veillonellaceae*, 17. *Acidaminococcaceae*; Proteobacteria: 1. *Sutterellaceae*, 2. *Enterobacteriaceae*, 3. Unknown *Rhodospirillales*.

B. Sample-specific distribution of single faecal OTUs according to their relative abundance (low, 0.1%–1.0%; high, >1.0%) and colour-coded according to the method used to isolate corresponding cultured species as in A.

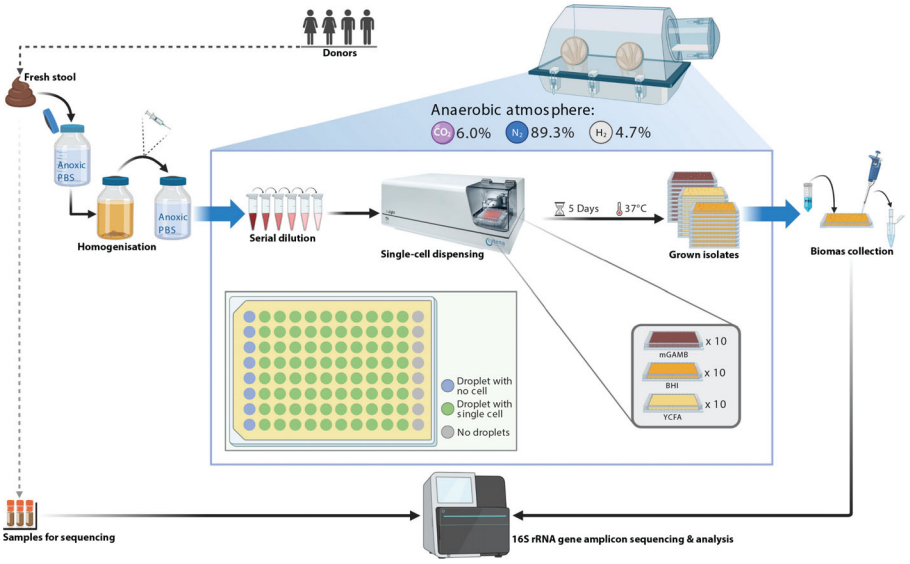

Fig. 3. Studying the cultured diversity after SCD isolation by bulk sequencing independent of individual strain processing for identification. Fresh faecal samples were collected from the same four donors as in the first experiment series. A total of 2400 single cells were dispensed per sample onto 10 plates for each culture medium (see their names in the scheme; same media as in the first experiments). Colonies from any given single plate were pooled (biomass collection; right-hand side of the scheme) and sequenced along with the original samples. Created with BioRender.

Interestingly, amplicon sequencing of the cultured biomass delivered 45 molecular species that were not captured by sequencing in the original faecal samples at relative abundances >0.1%, including seven novel taxa (Fig. 4D). The contribution of each sample to this cultivation-specific diversity varied, with 2–7 unique species per sample and another 17 species occurring in at least two samples. These species spanned all five dominant phyla, with a dominance of Firmicutes, Bacteroidota and Actinobacteriota. They generally occurred at an average relative abundance below 0.5% across 17 998 human intestinal amplicon datasets (Lagkouvardos *et al*., 2016a) (black bars in Fig. 4D). However, four of them were characterized by a prevalence ≥20%: *Prevotella copri*, *Bifidobacterium* sp., *Escherichia/Shigella* sp., and a novel genus within family *Bacteroidaceae* (<95% sequence identity to any species with a valid name).

Discussion

This work presents a SCD workflow that helps recover pure bacterial cultures from complex communities under anaerobic conditions. Considering the large fraction of yet uncultured bacteria in various ecosystems (Overmann *et al*., 2017; Thomas and Segata, 2019), this workflow can serve as a basis for accelerating the recovery of novel taxa and for establishing sample-specific collections of isolates. Whilst it can be applied to all types of input microbial communities, we focussed on the gut microbiota due to our own interest in this ecosystem and its importance for human health (O'Keefe, 2016; Sonnenburg and Sonnenburg, 2019; Visconti *et al*., 2019; Sepich-Poore *et al*., 2021).

Previous work has pioneered the establishment of individualized strain collections from the human gut by dilution-to-extinction (Goodman *et al*., 2011). Such an approach is limited by the stochastic recovery of clonal cultures. The application of image-based selection employed here improves the selection rate of droplets containing single cells and is thus more deterministic, reducing the risk of mixed cultures (Riba *et al*., 2016). Moreover, post-process quality control (QC) of the droplets can be conducted via analysis of the microscopic images stored for each dispensed droplet. Although the deterministic power of the SCD approach was not fully exploited in the present work due to the limited number of single isolates eventually identified, this will be addressed in future studies. Although the dispensing strategy presented here has been used in combination with single-cell genomics very recently (Wiegand *et al*., 2021), this is the first application of this technology to recover live anaerobic bacteria in a high-throughput manner.

The 'culturomics' approach proposed by others (Lagier *et al*., 2016) has helped to recover new bacterial diversity from the human gut. However, many of the bacteria isolated by this approach remain undescribed, and the throughput of the method relies solely on parallelising manual work, including tedious rounds of re-streaking. Using SCD, no plateau in the recovered species diversity was reached at a depth of 800 single cells from one diluted faecal slurry on a single culture medium. This can easily be up-scaled without a substantial increase in hands-on time at the beginning of the workflow by either dispensing more single cells on the same culture

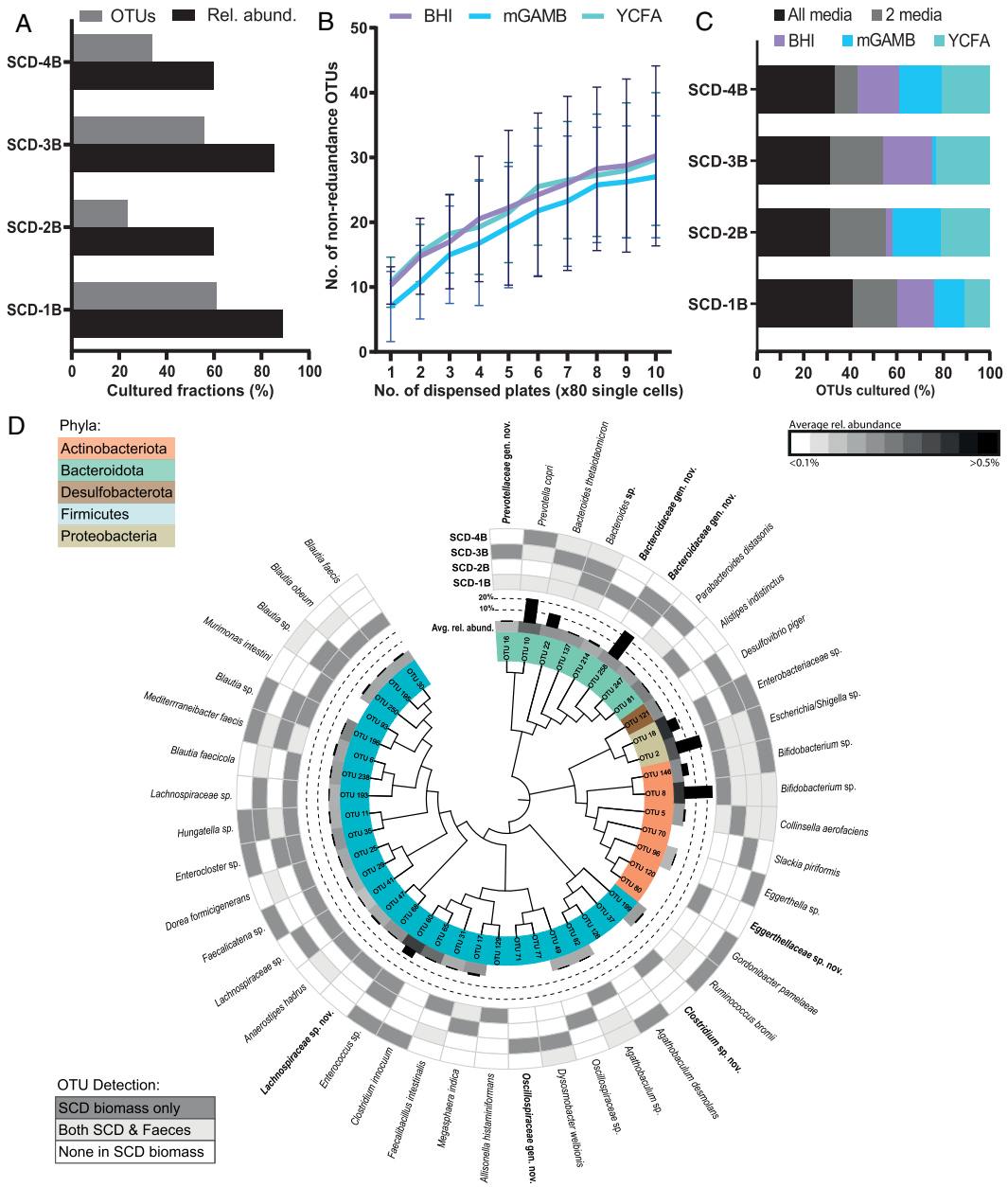

Fig. 4. Output of SCD cultivation without selection and identification of individual strains.
A. Cultured fraction per sample shown either as number of molecular species (grey bars) or corresponding cumulative relative abundance (black).
B. Cumulative number of OTUs captured as a function of the dispensing depth (number of plates per culture medium on the *x*-axis; 80 dispensed single cells per plate). The curves represent mean values, and the variations are shown as SD ($n = 4$ samples).
C. OTUs captured by cultivation based on the culture medium used.
D. Phylogenetic tree of all species captured by SCD but not detected in the original (directly sequenced) faecal samples. The colour-coded phylum classification is according to SILVA release 128 (Pruesse *et al*., 2012). The grey gradient in the first outer ring represents the average relative abundance of the corresponding OTUs in positive human intestinal samples out of 17 988 (positive = with a minimum relative abundance ≥0.1%). The prevalence of positive samples is represented by the black bars. Dark grey squares across the four samples indicate detection of the corresponding OTU in the given SCD biomass only (i.e. cell biomass collected from the SCD plate); light grey squares indicate detection of the corresponding OTU in both the SCD biomass and direct sequencing of the original stool sample; white squares indicate that the OTU was not detected in the given SCD biomass, but it may be in the corresponding faecal sample. OTUs were identified using EzBioCloud (Yoon *et al*., 2017b), based on a 97% and 95% sequence identity cut-off for species- and genus-level delineation respectively. Potential novel taxa are shown in bold.

medium, expanding the array of culture media for incubation, or using different pre-cultures of the same faecal slurry in different media to expand the variety of input community compositions prior to dispensing. This approach also eliminates the need to constantly monitor for the appearance of colonies after plating, which could be a problem when culturing on a large scale using the classical plate-based approach. Given that we have recently published a bioinformatic procedure that can speed up the description of novel taxa (Hitch *et al*., 2021b), the current bottleneck of the isolation workflow is the time required to identify all single isolates, which will also need to be automated in the future.

Several state-of-the-art cultivation projects have boosted the renewed interest in studying gut microbiomes by cultivation (Browne *et al*., 2016; Lagier *et al*., 2016; Seshadri *et al*., 2018; Zehavi *et al*., 2018; Forster *et al*., 2019; Zou *et al*., 2019; Groussin *et al*., 2021; Zenner *et al*., 2021; Beresford-Jones *et al*., 2022). Whilst such projects allow capturing a substantial cultured diversity due to their large scale (high number of sample/media combinations), the isolation depth per sample remains relatively low (Groussin *et al*., 2021). The primary aim of the current cultivation project was to benchmark a novel workflow rather than to perform large-scale work. Increasing the depth of SCD will soon allow establishing personalized strain collections. The collection of human gut isolates created (https://github.com/ClavelLab/SCD) contains 21 novel taxa that have been fully described taxonomically and helps further elucidating bacterial diversity within the human gut.

In conclusion, microbiome research is currently hampered by the lack of isolates from many microbial ecosystems. The renewed interest in cultivation emphasizes the need to develop bacterial isolation workflows with higher throughput. The present work represents a first step towards this goal.

Experimental procedures

Faecal sample collection and processing

Stool was collected from four healthy adults (26–36 years of age; two females, two males) residing in Aachen (Germany) and the surrounding area. Sample collection was approved by the Independent Ethics Committee of the RWTH University Hospital (file no. EK 10/20). Each participant received a stool collection kit consisting of a plastic bucket (0.5 L), a sterile plastic bag and an oxygen scavenger sachet (BD Biosciences; ref. 260683). Donors were instructed to defecate directly into the bucket lined with the plastic bag, which was then folded. The oxygen scavenger was placed in the bucket, which was then tightly sealed to minimize the exposure to oxygen. All samples were then processed under a laminar flow cabinet within 2 h after defecation. To prevent high amount of oxygen mixed into diluted sample, a two-step procedure was followed as presented previously (Wylensek *et al*., 2020). After manual homogenisation of the stool by kneading the plastic bag, approx. 5 g of faecal material was transferred into a 100 ml Schott bottle containing 50 ml anaerobic PBS supplemented with peptone (0.05% wt./vol.), L-cysteine (0.05% wt./vol.) and dithiothreitol (DTT) (0.02% wt./vol.) to obtain a ca. 10-fold dilution (wt./vol.). The bottle was flushed with an anaerobic gas mixture (89.3% N_2, 6% CO_2, 4.7 H_2) prior to closing the lid. After vigorous manual shaking to re-suspend the sample and sterilization of the rubber stopper by flaming twice with alcohol, 5 ml of the faecal slurry was rapidly transferred into another Schott bottle containing 45 ml of the same anaerobic medium using a syringe (flushed previously with the anaerobic gas mixture). These ca. 100-fold anaerobic faecal dilutions were then brought into the anaerobic workstation (MBraun GmbH, Germany) and further diluted with filter-sterilized anoxic PBS to a dilution of 10^{-8}. The PBS was filtered (0.2 μm) to remove any debris that may interfere with the camera detection or clog the nozzle of the SCD cartridge.

For 16S rRNA gene amplicon profiling of the samples, approx. 2 × 0.5 cm^3 of each stool sample was collected into a 2 ml screw-cap tube containing 500 mg autoclaved zirconia beads (Carl Roth, Germany; 0.1 mm) and 600 μl Stool DNA Stabilizer (Invitek Molecular, Germany). The tube was vortexed until obtaining a homogeneous suspension that was then frozen immediately on dry ice and stored at −80°C until use.

Culture media

All media compositions are given per litre of distilled water. The reducing agents L-cysteine and DTT were dissolved in 4.0 ml of distilled water and filter-sterilized before addition into the agar media after autoclaving (ca. 55°C) to reach a final concentration of 0.05% (wt./vol.) and 0.02% (wt./vol.) respectively. For the SCD approach, the agar medium was poured into classical 96-well plates or 1-well rectangular plates (Nunc™ OmniTray, Cat. No.: 140156). The following three media were applied:

Brain–Heart Infusion (BHI) Agar: BHI base (Oxoid, ref. CM1135), 37.0 g; agar, 15.0 g; resazurin, 1.0 mg.

Modified Gifu Anaerobic Media Agar with Blood (mGAMB): mGAM base (HyServe, ref. 05433), 41.7 g; agar, 15.0 g; defibrinated sheep blood, 50 ml. The blood

was added after sterilization when the agar medium reached ca. 55°C.

Yeast Casitone Fatty Acids (YCFA) Agar: DSMZ Medium 1611; Agar, 15.0 g; only DTT was added as reducing agent after autoclaving, as L-cysteine is already listed in the medium recipe.

All plates and dishes were prepared under sterile condition in a laminar flow cabinet. For 96-well plates, hot liquid agar was pipetted directly into the wells (250 µl per well) using a multichannel pipette. They were stored at 4°C for up to 21 days and brought into the anaerobic environment at room temperature 48 h prior to use.

Single-cell dispensing

General description. The single-cell dispenser b.sight (Cytena GmbH, Germany) consists of a microfluidic system, which constantly produces microdroplets of the target suspension, combined with optical components for microscopic picture acquisition and real-time analysis. After ejection from the nozzle of the cartridge filled with the diluted faecal slurry, the micro-droplets are either discarded by vacuum if containing no or multiple cells or dispensed onto the agar medium placed below in case a single cell was detected. This system was shown to efficiently separate cultures of both fluorescent and wild-type bacteria, including mixture of *Escherichia coli* and *Enterococcus faecalis* (Riba *et al*., 2016). It was placed into an anaerobic chamber (PlastLabs, USA) during the entire project. To control for sterility, agar plates containing each of the three culture media mentioned above remained open in the anaerobic workstation throughout the entire processing time. These control plates were incubated for 7 days and checked for colony formation, which never occurred.

Experiment type 1. The aim of this set of experiments was to assess process efficacy and output of the SCD approach when compared with CAP isolation (see description in the next section) (Fig. 1A). Approx. 70 µl of the 10^{-5} faecal dilution was pipetted into the SCD cartridge. Prior to dispensing, the droplet QC check was performed as per the manufacturer's instructions to ensure a reproducible and robust droplet generation (i.e. droplets were produced, and no multiple droplets were ejected at the same time) and that the vacuum suction of empty or multiple-cell droplets is functioning. This QC check is part of the SCD program and is performed prior to starting any experiment or after replacing the cartridge. It uses stroboscopic droplet imaging to observe the generated droplets live and in flight. The camera setting allows the re-positioning and focusing of droplets to the centre of the camera's field view. The droplet generation can be calibrated by adjusting the speed (droplet velocity) and length (droplet volume) of the displacement of the piston onto the silicon chip. Bacterial density was visually controlled and re-adjusted by adding higher or lower dilutions of the same sample whenever required. The agar plate was loaded, and cells were dispensed according to the dispensing template shown in Fig. 1A, where 80 droplets containing every single cell were dispensed per plate, the other 16 wells consisting of negative and positive controls. After dispensing, the plates ($n = 10$ per culture medium) were incubated at 37°C in the anaerobic chamber. Seven to eight colonies were picked for each medium at each time point (3 and 5 days of incubation), summing to 15 colonies per medium and 45 colonies per sample for identification. All SCD-derived cultures were pure, i.e. none of the picked colonies required re-streaking. Colony counts were determined at the end of the experiment (i.e. day 5). Calculation of hands-on time was done in 5-min intervals, whereas incubation was calculated in a 24-h interval.

Experiment type 2. The aim of this second set of experiments was to determine sample-specific cultured fractions based on the SCD approach independent of the single-strain identification steps, i.e. by sequencing bulk cultured biomass. The experimental setup is shown in Fig. 3. Dispensing was performed as described above. The total incubation time was 5 days to allow comparison with the other experiments. Plates were then removed from the anaerobic workstation and the biomass from all colonies on the same square agar dish was collected in PBS using a sterile L-loop. The biomass from large colonies was first partially removed using sterile single loops to reduce the risk of fast-growing bacteria to mask the presence of taxa that grew as smaller colonies. The pooled biomass was centrifuged (13 000*g*, 10 min), the pellet was re-suspended in 600 µl Stool DNA Stabilizer, and then frozen immediately on dry ice prior to storage at −80°C.

Classical agar plate cultivation

The starting material for isolation was the same faecal dilution series as used for the SCD experiments described above. Plating was performed in parallel. A volume of 50 µl (dilution 10^{-4} and 10^{-5}), 70 µl (10^{-6}), or 100 µl (10^{-7} and 10^{-8}) was pipetted onto each BHI, mGAM Blood and YCFA agar plates followed by spreading using single-use, sterile L-spatula. The plates were incubated for 5 days at 37°C. As in the SCD approach, seven to eight colonies were picked for each medium at each time point (3 and 5 days of incubation), with the aim of maximizing diversity by collecting colonies showing different morphologies whenever possible. This resulted in a number of 15 colonies per medium and 45 colonies in

total per sample for identification. The selected colonies were re-streaked three times to ensure purity prior to identification. All processing steps and incubation occurred under anaerobic conditions, including the same negative controls as in the SCD experiments. Note that the overall time required for the CAP procedure depends on the number of re-streaking rounds. Hence, the number of three rounds in these experiments increased the time accordingly. However, the main priority of this work is to provide a solid foundation for either the description of novel taxa (hence purity is essential for deposition in national collections) or for downstream work in gnotobiotic models. Moreover, the long-term goal of this project is to provide a framework for studying large numbers of isolates using an automated workflow. With such large numbers, going back to potentially contaminated isolates in case only one re-streaking is performed as standard would eventually lead to more work and confusion. These reasons have driven the choice of three times re-streaking, which is also commonly reported in textbooks and deemed 'good practice'.

Strain identification

All isolates were first identified using a MALDI-Biotyper (Bruker Daltonik) equipped with the up-to-date public reference database at the time of analysis (as per September 2020). Cell biomass from each colony was directly applied onto the MALDI target. The MALDI matrix (α-Cyano-4-hydroxycinnamic acid) was added (1 μl) on the dried samples, which were then measured following the manufacturer's instructions. For isolates not identifiable at the species level by MALDI (score <1.7), the 16S rRNA gene was sequenced. The gene was amplified by colony PCR using the primers 27f and 1492r (Greuter *et al*., 2016). PCR products were checked by gel electrophoresis and cleaned using the QIAquick PCR Purification Kit (Qiagen, Germany). Sanger sequencing was performed at Eurofins Genomics (Ebersberg, Germany) or Microsynth Seqlab (Göttingen, Germany) using the primers 27f, 1492r, 338r and 785r (Clavel *et al*., 2013; Wylensek *et al*., 2020). Raw sequences were checked and corrected manually with help of the electropherograms and contigs were generated to obtain near full-length gene sequences. Most closely related species with a valid name were identified using EzBioCloud (Yoon *et al*., 2017b). A cut-off value of 98.7% was used to delineate known from novel species.

Strain preservation

All pure isolates were eventually cultured in Hungate tubes containing liquid medium for easy handling and biomass generation. The bacteria that could not be maintained in culture during re-streaking (CAP approach) or in Hungate tubes (CAP and SCD approach) were referred to as 'stopped growing'. For long-term storage, fresh liquid cultures were mixed with anoxic Wilkins-Chalgren Anaerobe (WCA) medium containing 40% glycerol (Sigma-Aldrich, Germany) in a 1:1 ratio, yielding a final concentration of 20% glycerol. This mixture was prepared under sterile condition in a laminar flow cabinet using a syringe and aliquoted into several screw-cap tubes. For a few isolates that could not be grown in liquid medium, an agar plate culture kept in the anaerobic workstation was flooded with anoxic WCA containing 20% glycerol, mixed with a sterile L-spatula, and the mixture was aliquoted into screw-cap tubes. All the tubes were then frozen immediately on dry ice and stored at −80°C. All isolates corresponding to novel taxa and described in the present study were submitted to the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany).

Genome sequencing and processing

The isolates were revived from frozen glycerol stocks and grown on liquid or agar media to obtain bacterial biomass. The DNA was isolated using a modified version of the method by Godon *et al*. (1997). Cell lysis included bead-beating in the presence of DNase inhibitor and detergent followed by purification on NucleoSpin gDNA Clean-up columns (Macherey-Nagel, Germany). DNA integrity was checked by gel electrophoresis and concentration was measured using Qubit fluorometer (Thermo Fischer Scientific, USA).

Most of the DNA libraries were prepared with the NEBNext Ultra II FS DNA Library Prep Kit for Illumina (NEB, USA) according to the manufacturer's protocol using 300 ng of DNA. The time used for enzymatic shearing to ca. 200 bp was 30 min. The PCR enrichment of adaptor-ligated DNA was conducted with five cycles and NEBNext Multiplex Oligos for Illumina (NEB) for paired-end barcoding. For size selection and clean-up of adaptor-ligated DNA, AMPure beads (Beckman Coulter, USA) were used. Quality check (Bioanalyzer System, Agilent Technologies, USA) and DNA quantification (Quantus, Promega, USA) of resulting libraries were conducted at the IZKF Core Facility (UKA, RWTH Aachen University), as was the sequencing on a NextSeq500 (Illumina, USA) with a NextSeq500 Mid Output Kit v2.5 (300 Cycles). The DNA libraries for isolates CLA-AA-H200 and CLA-AA-H284 were prepared as described above, but the time used for enzymatic shearing to approximate 450 bp was 10 min and paired-end sequencing was conducted on a MiSeq (Illumina) with a MiSeq reagent kit v3 (600 cycles). The DNA libraries for isolates CLA-AA-H204 were prepared using mechanical shearing of 300–500 ng of DNA

on a Covaris M220 to 400–500 bp (50s; peak incident power, 75 W; duty factor, 10%; cycles per burst, 200). For PCR enrichment of adaptor-ligated DNA, 6 cycles and NEBNext Multiplex Oligos for Illumina for single barcoding were used. DNA concentration was measured using a Qubit fluorometer (Thermo Fisher Scientific, USA), before paired-end sequencing on a MiSeq (Illumina) using the v3 chemistry (600 cycles).

Raw reads were quality-filtered and adapters, as well as phiX reads, were removed using Trimmomatic v0.39 (Bolger *et al*., 2014) and bbduk (Bushnell, 2014). Assemblies were obtained using SPAdes v3.13.1 (Bankevich *et al*., 2012) followed by quality check using CheckM (Parks *et al*., 2015). Genomes were deemed to be of high quality when >95% complete and containing ≤5% contamination.

Scanning electron microscopy

SEM was performed for all novel taxa at the Electron Microscopy Facility of the RWTH University Hospital. Bacterial biomass was fixed with 3% (vol./vol.) glutaraldehyde (Agar Scientific, UK) in 0.1 M Sorensen's phosphate for 1 h, then washed in phosphate buffer for 15 min, and dehydrated by incubating in an ascending ethanol series (30%, 50%, 70%, 90% and 100%) for 10 min each and the last step thrice. Samples were then critical point dried in liquid CO_2 (Polaron, GaLa Instrumente, Germany) and sputter coated (Sputter Coater EM SCD500, Leica, Germany) with a 10-nm gold/palladium layer. They were subsequently analysed using an environmental scanning electron microscope (ESEM XL-30 FEG, FEI Company, Netherlands) with a 10-kV acceleration voltage in a high vacuum environment.

High-throughput 16S rRNA gene amplicon sequencing

Metagenomic DNA was isolated from the frozen stool aliquots or biomass collected from agar plates using a modified version of the method by Godon *et al*. (1997) as briefly described for genome sequencing. DNA integrity was checked by gel electrophoresis and concentration was measured using Nanodrop (Thermo Fischer Scientific). The V3–V4 regions of 16S rRNA genes were amplified in a two-step approach (15 + 10 cycles) (Berry *et al*., 2011) using the primers 341F and 785R (Klindworth *et al*., 2013) following a combinatorial dual indexing strategy. Libraries were purified using AMPure XP magnetic beads (Beckman-Coulter, Germany), pooled in equimolar amounts, and sequenced in paired-end mode using the v3 chemistry (600 cycles) on an Illumina MiSeq according to the manufacturer's instructions. The platform was semi-automated (Biomek4000 pipetting robot, Beckman Coulter, Germany) to increase reproducibility and the workflow systematically included two negative controls: a DNA-extraction control, i.e. sample-free DNA-stabilization solution and a PCR blank, i.e. PCR-grade water as template, for every 46 samples sequenced.

Amplicon sequence analysis

Raw sequencing reads were processed using IMNGS (www.imngs.org) (Lagkouvardos *et al*., 2016a), a platform based on UPARSE (Edgar, 2013). A sequence identity threshold of 97% was used for clustering sequences into OTUs. Other parameters were: barcode mismatch tolerated, 1; min. trim quality score, 3; max. expected errors, 3; max. mismatch during merging of read, 50; min. percentage identity of alignment during merge, 70%; min. read length, 350; max. read length, 500; no. of nucleotides trimmed at each the 5′- and 3′-end, 10. Only OTUs that occurred at a relative abundance ≥0.1% in at least one sample were kept for further processing (Reitmeier *et al*., 2021). OTUs were taxonomically classified using SILVA (Pruesse *et al*., 2012). Further data processing (diversity and composition analyses) was done in R using Rhea (Lagkouvardos *et al*., 2017a), including the use of normalized relative abundances to allow comparisons. The 16S rRNA gene sequences of the isolates were matched to OTUs using blastn (*E*-value <1e−25, 97% identity, 80% query coverage). The species dendrogram in Fig. 2A was created in R using Vegan, Ape and Taxize based on their taxonomic classification (SILVA), and further edited using iTOL version 6.3 (Letunic and Bork, 2021). For the tree in Fig. 4D, 16S rRNA gene sequences were aligned using the SILVA aligner (Pruesse *et al*., 2012), the alignment was curated manually, and phylogeny was inferred in MEGA 7 (Kumar *et al*., 2016) using the maximum likelihood method.

Statistics

Unless otherwise specified, values are present as mean ± SD. Results from the SCD and CAP approaches were compared statistically by calculating *p*-values using the Mann–Whitney *U* test using GraphPad Prism version 9.1.2 (GraphPad Software, USA).

Taxonomic descriptions

Isolates were assumed to represent novel taxa based on <98.7% 16S rRNA gene sequence identity to bacteria with valid names using EzBioCloud (Yoon *et al*., 2017b). Novelty was further confirmed by comparison to the LPSN database (01/01/2022). Bacterial names were required to be both correct and valid for the matching isolate to be termed a member of that species, else the

isolate was deemed novel. The genomes of novel candidates were sequenced followed by taxonomic, ecological and functional analyses using Protologger (www.protologger.de) (Hitch *et al*., 2021b). All Protologger output files are available in the project data repository at https://github.com/ClavelLab/SCD. With this tool, the 16S rRNA gene sequence was chimaera checked (Usearch v5.2.32, Edgar *et al*., 2011) followed by closest relatives identification using SILVA-LTP (Pruesse *et al*., 2007; Yilmaz *et al*., 2014) and ecological analysis based on IMNGS (Lagkouvardos *et al*., 2016a). Genomes were used to determine taxonomy at multiple levels: The Genome Taxonomy Database Toolkit (GTDB-Tk) (Parks *et al*., 2018; Chaumeil *et al*., 2019), average nucleotide identity (ANI) (Jain *et al*., 2018), percentage of conserved proteins (POCP) (Qin *et al*., 2014) and mol.% G + C content (Meier-Kolthoff *et al*., 2014). The analysis of functional features included: CDS and CRISPR (Seemann, 2014), antibiotic resistance (McArthur *et al*., 2013), carbohydrate degradation (Lombard *et al*., 2014) and KEGG pathways (Ogata *et al*., 1999). The ecological analysis based on 16S rRNA gene sequences mentioned above was complemented by genome comparison to metagenome-assembled genomes using MASH (Ondov *et al*., 2016). The following thresholds were considered to assign taxonomies: <98.7%, <94.5% and <86.5% based on 16S rRNA gene sequence identities as indications for novel species, genera and families respectively (Yarza *et al*., 2014); ANI values <95% and genome-based differences in G + C content of DNA >1% (Meier-Kolthoff *et al*., 2014) to separate species; POCP values <50% for distinct genera (Qin *et al*., 2014). The classification of isolates based on genomes was also evaluated by constructing phylogenomic trees to either confirm or deny novelty and genus or family delineation. In addition, manual POCP analysis (Qin *et al*., 2014) and Genome-to-Genome Distance Calculator 3.0 (GGDC, http://ggdc.dsmz.de/ggdc.php#) (Meier-Kolthoff *et al*., 2014) with a cut-off of 70% for species-level were also used along with online ANI Calculator (https://www.ezbiocloud.net/tools/ani) (Yoon *et al*., 2017a) for delineation of certain species. Manual genome annotation (e.g. butyrate biosynthesis) was also performed on selected species using Prokka (Seemann, 2014), version 1.14.6. Scanning electron micrographs of all the isolates representing novel taxa are available online (Supplementary File S5; https://github.com/ClavelLab/SCD).

Description of *Bacteroides hominis* sp. nov. *Bacteroides hominis* (ho'mi.nis. L. gen. masc. n. *hominis*, of a human being, referring to the human gut habitat).

This isolate shared 98.41% 16S rRNA gene sequence similarity to *Bacteroides fragilis*. The genome analysis also identified *B. fragilis* as the closest relative but confirmed the status of novel species with an ANI and GGDC value of 87.4% and 33.1% respectively. The POCP value of 78.5% to *B. fragilis*, the type species of this genus, confirmed that the isolate belongs to the genus *Bacteroides*. The isolate was found to be the same species as 'Bacteroides hominis' (Liu *et al*., 2021), with ANI and GGDC values of 98.1% and 84.8% respectively. However, this species has not yet been validated. Cells are rods with generally 1.0–2.5 μm in length when grown on BHI under anaerobic conditions. Genome analysis predicted the ability to utilize starch, sulfide and L-serine. The genes for production of acetate, propionate, L-glutamate and riboflavin (vitamin B2) were also detected. Several antibiotic-resistant genes were identified: Erm 23S ribosomal RNA methyltransferase (ARO:3000560), CcrA beta-lactamase (ARO:3004200), tetracycline inactivation enzyme (ARO:3000036) and tetracycline-resistant ribosomal protection protein (ARO:0000002). The G + C content is 43.6 mol.%, close to *B. fragilis* (43.1 mol.%). The type strain, CLA-AA-H207^T (=DSM 112683^T) was most prevalent in human gut microbiota (71.1% of 1000 samples positive), followed by mouse and pig gut microbiota with 49.6% and 46.8% respectively. It was isolated from the faeces of a healthy 26-year-old woman.

Description of *Blautia fusiformis* sp. nov. *Blautia fusiformis* (fu.si.for'mis. L. masc. n. *fusus*, spindle; L. fem. n. *forma*, shape, form; N.L. fem. adj. *fusiformis*, spindle-shaped, pertaining to the cell morphology).

The isolates CLA-AA-H217 and CLA-AA-H275 had the highest 16S rRNA genes sequence identity to multiple *Blautia* species, particularly *Blautia obeum* (96.80%–97.07%), indicating its position within the genus *Blautia*. The assignment based on POCP value (82.0% to *Blautia obeum*) and classification according to GTDB-Tk also support the genus assignment. No validly named species with a sequenced genome within the GTDB-Tk database was identified with ANI value above 95% (77.15%–77.48% to *Blautia intestinalis*), thus confirming the isolate as a novel species within the genus *Blautia*. The isolate was found to be the same species as 'Blautia massiliensis' (Durand *et al*., 2017b; Liu *et al*., 2021), with ANI and GGDC values of 97.9%–98.3% and 83.2%–86.7% respectively. However, this species name has never been validated. The species grows on modified Gifu Anaerobic Medium under anaerobic conditions, forming short rod-shaped cells with slightly pointy end (length 1.5–2.5 μm). Genome analysis identified the presence of genes for utilization of glucose and starch, and production of acetate, propionate, L-glutamate, cobalamin (vitamin B12), folate and riboflavin (vitamin B2). No antibiotic resistance genes were detected. The G + C content is 44.1 mol.%, whereas that of *B. obeum* is 41.6 mol.% and

B. coccoides (type species) 45.6 mol.%. The species was most prevalent in human gut microbiota (69.3%–70.5% of 1000 samples positive), followed by pig gut microbiota (55.0%–56.3%), wastewater (46.6%–53.5%), activated sludge (39.8%–52.4%) and chicken gut microbiota (50.3%–50.4%). The type strain, CLA-AA-H217^T (=DSM 112726^T) was isolated from the faeces of a healthy 26-year-old woman. The other strain within this species, CLA-AA-H275 (=DSM 113286), was isolated from the faeces of a healthy 30-year-old man.

Description of *Coprococcus hominis* sp. nov. *Coprococcus hominis* (ho'mi.nis. L. gen. masc. n. *hominis*, of a human being, referring to the human gut habitat).

The closest relatives to the strain are *Coprococcus eutactus* (94.23%), the type species of the genus *Coprococcus*, and *Eubacterium ruminantium* (94.11%) based on the 16S rRNA gene sequence similarity. None of the closely related species identified had an ANI value above 95% (*C. eutactus*, 80.9%; *E. ruminantium*, 77.6%). GTDB-Tk assigned the species to genus 'CAG-127' within the family *Lachnospiraceae*. The closest relative based on the genome tree is *E. ruminantium* (followed by *C. eutactus*). However, the POCP value was 51.4% to *C. eutactus*, while it was 21.7% to *Eubacterium limosum*, the type species of the genus *Eubacterium*. Therefore, the isolate was considered to represent a novel species within the genus *Coprococcus*. The isolate was found to be the same species as 'Wujia chipingensis' (Liu *et al*., 2021), with ANI and GGDC values of 98.3% and 84.4% respectively. However, this species has not yet been validated. Cells are short rods (0.7–1.0 μm in length) (Supplementary Files) that grow well on modified Gifu Anaerobic Medium with 5% Sheep blood under anaerobic conditions. Genome analysis predicted the ability to utilize starch and produce acetate, propionate, L-glutamate, cobalamin (vitamin B12) and folate. Antibiotic resistance may be present due to the detection of *vanU* (ARO:3000575) and tetracycline-resistance ribosomal protection protein (ARO:0000002). The G + C content is 44.2 mol.%, which varies by more than 1% to that of *C. eutactus* (43.1 mol.%). The type strain, CLA-AA-H212^T (=DSM 112732^T), was most prevalent in wastewater (37.3% of 1000 samples positive), followed by pig gut microbiota (33.8%) and human gut microbiota (23.9%). The bacterium was isolated from the faeces of a healthy 26-year-old woman.

Description of *Faecalibacterium hominis* sp. nov. *Faecalibacterium hominis* (ho'mi.nis. L. gen. masc. n. *hominis*, of a human being, referring to the human gut habitat).

The strains CLA-AA-H223, CLA-AA-H254 and CLA-AA-H283 had the highest 16S rRNA gene sequence similarities to *Faecalibacterium prausnitzii* (98.04%–98.38%). The highest ANI value was to *F. prausnitzii* (85.6%–85.8%). GTDB-Tk placed the genome under the genus *Faecalibacterium*, and species 'Faecalibacterium_prausnitzii_C'. The highest POCP values were 71.2%–74.1% to *F. prausnitzii*. The isolates were also grouped together with *F. prausnitzii* based on the genome tree analysis. These analyses confirmed the classification of the isolates as a novel species within the genus *Faecalibacterium*. The isolates were found to represent the same species as 'Faecalibacterium hominis' (Liu *et al*., 2021), with ANI and GGDC values of 96.5%–96.9% and 72.5%–73.7% respectively. However, this species has not yet been validated. Cells are long rods to string like on modified Gifu Anaerobic Medium under anaerobic conditions. Genome analysis predicted the ability to utilize glucose, arbutin, salicin, starch, sulfide and L-serine in all three strains. The genes for utilization of trehalose were detected in both CLA-AA-H223 and CLA-AA-H254, whereas the genes for cellobiose utilization were only found in CLA-AA-H223. The genes for production of acetate, propionate, L-glutamate, L-cysteine, cobalamin (vitamin B12) and riboflavin (vitamin B2) were present in all strains. As butyrate biosynthesis was not predicted, manual examination of the Prokka annotations identified genes assigned as 3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157) and butyryl-CoA:acetate CoA-transferase (EC 2.8.3.-). Antibiotic resistance genes were only detected in the genome of strain CLA-AA-H254: Erm 23S ribosomal RNA methyltransferase (ARO:3000560). The molecular G + C content of DNA is 56.4–56.9 mol.%. The species was most prevalent in human gut microbiota (77.3%–79.8% of 1000 samples positive), followed by wastewater (60.9%–68.1%), pig gut microbiota (57.4%–67.4%) and activated sludge (60.9%–65.0%). The type strain, CLA-AA-H223^T (=DSM 113002^T), was isolated from the faeces of a healthy 26-year-old woman. The other strain of this species, CLA-AA-H254 (=DSM 113716) and CLA-AA-H283 (=DSM 113416) were isolated from the faeces of a healthy 36-year-old woman and 28-year-old man respectively.

Description of *Fusicatenibacter faecihominis* sp. nov. *Fusicatenibacter faecihominis* (fae.ci.ho'mi.nis. L. fem. n. *faex faecis*, the dregs, *faeces*; L. gen. n. *hominis*, of a human being; N.L. gen. n. *faecihominis*, of human faeces).

Based on 16S rRNA gene sequences, the closest taxonomic neighbours of the isolate are species within the family *Lachnospiraceae*, such as *Clostridium nexile* (93.59%) and *Ruminococcus lactaris* (93.51%). The closest genome in terms of ANI value was *Coprococcus comes* (83.0%). While GTDB-Tk placed the genome under genus 'GCA-900066135' (family

Lachnospiraceae), the highest POCP value was 54.5% to *Fusicatenibacter saccharivorans*, the type species of this genus. Furthermore, genome tree placement next to this species confirmed the novel species status of the isolate within the genus *Fusicatenibacter*. Cells are spindle shaped (generally 1–2.5 μm in length) on YCFA medium under anaerobic conditions. Genome analysis predicted the ability to utilize arbutin, salicin and starch. The genes for production of acetate, propionate, L-glutamate, cobalamin (vitamin B12) and folate were also identified. The detection of tetracycline-resistant ribosomal protection protein (ARO:0000002) indicates antibiotic resistance. The G + C content of DNA is 48.4 mol.%. The type strain, CLA-AA-H277^T (=DSM 113288^T), was most prevalent in human gut microbiota (40.0% of 1000 samples positive), followed by pig gut microbiota (37.4%) and wastewater (37.0%). It was isolated from the faeces of a healthy 30-year-old man.

Description of *Intestinimonas aquisgranensis* sp. nov. *Intestinimonas aquisgranensis* (a.quis.gra.nen'sis. M.L. fem. adj. *aquisgranensis*, pertaining to the city of Aachen in Germany, where the bacterium was isolated).

The isolate shared the highest 16S rRNA gene sequence similarity to *Intestinimonas butyriciproducens* (96.84%) and *Lawsonibacter asaccharolyticus* (96.13%). The highest ANI value was to *Pseudoflavonifractor capillosus* with 83.7%, confirming the novel status of the isolate. The highest POCP value (59.9%) was to *Intestinimonas butyriproducens*, the type species of this genus, followed by *Pseudoflavonifractor capillosus* (59.7%). However, the genomic tree analysis placed the strain closer to *I. butyriproducens* than to *P. capillosus*. GTDB-TK also assigned the genome to the genus *Intestinimonas*, further supporting the classification of the isolate as a novel species within this genus. The isolate was found to be the same species as 'Intestinimonas timonensis' (Durand *et al*., 2017a), with ANI value of 96.23%. However, this species name has never been validated. Genome analysis predicted the ability to utilize starch, sulfide and L-serine, and to produce acetate, butyrate, propionate, L-cysteine and riboflavin (vitamin B2). The detection of tetracycline-resistant ribosomal protection protein (ARO:0000002) may indicate the antibiotic resistance potential of this strain. It grows under anaerobic conditions on BHI media and forms cells that are rod shaped with slightly pointed ends. The G + C content is 59.8 mol.%. The type strain, CLA-AA-H199^T (=DSM 112680^T), was most prevalent in mouse gut microbiota (27.1% of 1000 samples positive), followed by chicken gut microbiota (24.5%) and human gut microbiota (23.4%). It was isolated from the faeces of a healthy 28-year-old man.

Description of *Ruminococcus turbiniformis* sp. nov. *Ruminococcus turbiniformis* (tur.bi.ni.for'mis. L. masc. n. *turben*, spinning top; L. fem. n. *forma*, shape; N.L. masc. adj. *turbiniformis*, shaped like a spinning top, referring to the shape of the bacterium).

The isolate had the highest 16S rRNA gene sequence identity to *Ruminococcus torques* (96.46%). The highest ANI value was 79.40% to *Enterocloster lavalensis*, while the ANI value to *R. torques* was 78.41%. GTDB-TK assigned the genome to an unknown species within the genus *Faecalicatena*. The POCP value to *R. flavefaciens* (type species, 25.3%) and *Faecalicatena contorta* (type species, 44.1%) was below genus delineation value, yet the values to *R. torques* and *R. faecis* were 52.4% and 51.0% respectively. The genome tree analysis clustered the isolates together with *Ruminococcus* species, therefore confirmed the status as a novel species within this genus within family *Oscillospiraceae*. The bacterium formed spinning top like, coccobacilli cells with approx. 1.0 μm in length on YCFA medium under anaerobic conditions. Genome analysis predicted the ability to utilize glucose, starch, sulfide and L-serine. The genes for production of acetate, propionate, L-cysteine L-glutamate and folate were also detected. Antibiotic resistance genes related to tetracycline-resistant ribosomal protection protein (ARO:0000002), Erm 23S ribosomal RNA methyltransferase (ARO:3000560) and major facilitator superfamily antibiotic efflux pump (ARO:0010002) were also present. The G + C content is 49.2 mol.%. The type strain, CLA-AA-H200^T (=DSM 112681^T), was most prevalent in chicken gut microbiota (83.7% of 1000 samples positive), followed by human gut microbiota (39.7%) and pig gut microbiota (21.2%). It was isolated from the faeces of a healthy 28-year-old man.

Description of *Veillonella fallax* sp. nov. *Veillonella fallax* (fal'lax. L. fem. adj. *fallax*, deceitful, referring to the tricky nature of identification of this species using 16S rRNA gene sequencing).

Based on the 16S rRNA gene sequence, the isolate shared >99.0% sequence identity to *V. infantium*, *V. nakazawae*, *V. dispar* and *V. tobetsuensis*. However, genome analysis showed the closest relatedness to *V. dispar* with an ANI value of 94.99% and GGDC value of 57.50%, confirming the novel species status of this isolate. Furthermore, GTDB-Tk classified the isolate as a novel species within the genus *Veillonella*. The POCP value of 88.00% with the type species of the genus *Veillonella* (*V. parvula*) further supports this classification. Cells are coccoid with diameter of ca. 0.5 μm when grown on YCFA agar under anaerobic conditions. Genome analysis predicted the production of propionate, and biosynthesis of folate and riboflavin (vitamin B2). No antibiotic resistance genes were identified. The G + C content of the genome is 38.7 mol.%, close to *V. dispar* (38.9 mol.%). The type strain, CLA-AA-H247^T (=DSM 113117^T), was most

prevalent in human oral microbiota (95.7% of 1000 samples positive), followed by human lung microbiota (79.2%) and human gut microbiota (67.4%). It was isolated from the faeces of a healthy 30-year-old man.

Description of *Hominilimicola* gen. nov. *Hominilimicola* (Ho.mi.ni.li.mi'co.la. L. masc. n. *homo*, a human being; L. masc. n. *limus*, dung; L. masc./fem. suffix *-cola*, an inhabitant of; N.L. masc. n. *Hominilimicola*, a microbe from the faeces of humans).

The closest relatives based on 16S rRNA gene sequence similarity are species within the order *Eubacteriales* and family *Oscillospiraceae*, with the highest similarity to *Monoglobus pectinilyticus* (89.0%). The highest POCP comparison value to the closest known species was 38.0% (*M. pectinilyticus*). While GTDB-Tk placed the genome under the order 'Monoglobales' (not validly published) but unable to provide family-, genus- or species-level assignment, this genus is phylogenetically placed into the family *Oscillospiraceae* (phylum Firmicutes) based on genome tree analysis (Supplementary Fig. S1). It formed a unique distinct branch to the only closest cultured relative (*M. pectinilyticus*), confirming the novelty of this genus. The type species is *Hominilimicola fabiformis*.

Description of *Hominilimicola fabiformis* sp. nov. *Hominilimicola fabiformis* (fa.bi.for'mis. L. fem. n. *faba*, a bean; L. fem. n. *forma*, shape; N.L. masc. adj. *fabiformis*, bean-shaped, referring to the cell shape).

The species has all features of the genus. Cells were oval, mung-bean like (length 0.5–1.0 μm, width 0.5 μm) in BHI medium under anaerobic conditions. Genome analysis predicted the utilization of glucose and starch, and the production of acetate, propionate, L-glutamate, cobalamin (vitamin B12) and folate. Antibiotic resistance may be present due to the detection of tetracycline-resistant ribosomal protection proteins (ARO:0000002). The G + C content of the genome is 37.8 mol.%. The type strain, CLA-AA-H232^T (=DSM 113452^T), was most prevalent in wastewater (55.1% of 1000 samples positive), followed by human gut microbiota (50.4%) and activated sludge (40.6%). It was isolated from the faeces of a healthy 30-year-old man.

Description of *Hominicoprocola* gen. nov. *Hominicoprocola* (Ho.mi.ni.co.pro'co.la. L. masc. n. *homo*, a human being; Gr. fem. n. *kópros*, dung; N.L. suffix masc./fem. *-cola*, an inhabitant of; N.L. masc. n. *Hominicoprocola*, a microbe from the faeces of humans).

The genus shared the highest 16S rRNA gene sequence similarities to *Dysosmobacter welbionis* and *Oscillibacter ruminantium* with 93.48% and 93.31% respectively. None of the identified closest relatives had ANI values above 95%. The analyses based on POCP, genome tree and GTDB-Tk all classified the isolate as a novel genus within family *Oscillospiraceae*. The type species is *Hominicoprocola fusiformis*.

Description of *Hominicoprocola fusiformis* sp. nov. *Hominicoprocola fusiformis* (fu.si.for'mis. L. masc. n. *fusus*, spindle; L. fem. n. *forma*, shape, form; N.L. masc. adj. *fusiformis*, spindle-shaped, pertaining to the cell morphology).

The species has all features of the genus. The isolate formed spindle-shaped cells with average 1–2 μm length on YCFA medium under anaerobic conditions. Genome analysis predicted the ability to utilize starch and produce acetate, propionate and L-glutamate. The detection of lincosamide nucleotidyltransferase (ARO:3000221) within the genome suggests antibiotic resistance. The G + C content is 55.7 mol.%. The type strain, CLA-AA-H269^T (=DSM 113271^T), was most prevalent in pig gut microbiota (65.3% of 1000 samples positive), followed by human gut microbiota (49.7%) and wastewater (39.1%). It was isolated from the faeces of a healthy 36-year-old woman.

Description of *Brotaphodocola* gen. nov. *Brotaphodocola* (Brot.a.pho.do'co.la. Gr. masc. n. *brotós*, a mortal human; Gr. fem. n. *ápodos*, dung; L. masc./fem. suffix *-cola*, an inhabitant of; N.L. masc. n. *Brotaphodocola*, a microbe from the faeces of humans).

The closest species to this genus based on 16S rRNA gene sequence comparison was *Clostridium fessum* (94.25%) and *Lacrimispora amygdalina* (94.20%). GTDB-Tk assigned the genome as an unknown species within genus 'UBA9502' (family *Lachnospiraceae*). None of the close relatives had ANI values above 95% to the isolate. Whilst the genome tree analysis placed the isolate next to *Hungatella hathewayi*, the POCP value to this species was only 40.5%. A similar value (40.4%) was also calculated to *H. effluvii* (the type species of the genus *Hungatella*), confirming the novel genus status of this isolate within family *Lachnospiraceae*. Note that both *H. effluvii* and *C. fessum* are assigned within the family *Clostridiaceae* in LPSN. However, our genome tree and the recent publication by Haas and Blanchard (2020) placed *Hungatella* spp. within family *Lachnospiraceae*. The type species is *Brotaphodocola catenula*.

Description of *Brotaphodocola catenula* sp. nov. *Brotaphodocola catenula* (ca.te'nu.la. L. fem. dim. n. *catenula*, a small chain).

The species has all features of the genus. Cells are spindle shaped to rods with slightly pointy ends (length 2–3.5 μm) and tend to form chains when grown under anaerobic conditions on YCFA medium. Genome analysis predicted the ability to utilize glucose, arbutin, salicin, starch, sulphide and L-serine. The genes for production

of acetate, propionate, L-glutamate, L-cysteine, folate and riboflavin (vitamin B2) were also identified. Antibiotic resistance may be present due to detection of genes encoding for tetracycline-resistant ribosomal protection protein (ARO:0000002) and Erm 23S ribosomal RNA methyltransferase (ARO:3000560). The G + C content is 47.2 mol.%. The type strain, CLA-AA-H274^T (=DSM 113285^T), was most prevalent in chicken gut microbiota (67% of 1000 samples positive), followed by human gut microbiome (56.5%) and pig gut microbiota (50.7%). It was isolated from the faeces of a healthy 30-years-old man.

Description of *Hominenteromicrobium* gen. nov. *Hominenteromicrobium* (Ho.min.en.te.ro.mi.cro'bi.um. L. masc. n. *homo*, a human being; Gr. neut. n. *énteron*, the gut; L. neut. n. *microbium*, a microbe; N.L. neut. n. *Hominenteromicrobium*, a microbe from the intestines of humans).

The closest phylogenetic neighbours based on 16S rRNA gene sequence similarity are species within family *Oscillospiraceae*: *Clostridium jeddahense* (still assigned to family *Clostridiaceae* in LPSN), 93.87%; *Caproiciproducens galactitolivorans*, 93.68%; *Hydrogeniiclostidium mannosilyticum*, 93.25%. None of the validly named species was identified to have ANI values above 95%. GTDB-Tk assigned the genome to the genus 'UBA1417' (family 'Acutalibacteraceae'). The genome tree placed the isolate next to *Acutalibacter muris*. The POCP value of 46.5% to this species confirmed the novel genus status within family *Oscillospiraceae*. The type species is *Hominenteromicrobium mulieris*.

Description of *Hominenteromicrobium mulieris* sp. nov. *Hominenteromicrobium mulieris* (mu.li.e'ris. L. gen. n. *mulieris*, of a woman, referring to the gut of a woman where the bacterium was isolated).

The species has all features of the genus. Cells are spindle-shaped coccobacilli, with approx. 1.5 μm in length, which often form pairs when grown under anaerobic conditions on BHI medium. Genome analysis predicted the ability to utilize starch, cellulose, sulfide and L-serine. Genes for the production of propionate, L-cysteine and L-glutamate were also identified. No antibiotic resistance genes were detected. The G + C content is 52.7 mol.%. The type strain, CLA-AA-H250^T (=DSM 113252^T), was most prevalent in pig gut microbiota (63.2% of 1000 samples positive), followed by wastewater (55.1%) and chicken gut microbiota (52.0%). It was isolated from the faeces of a healthy 36-year-old woman.

Description of *Brotocaccenecus* gen. nov. *Brotocaccenecus* (Bro.to.cacc.en.e'cus. Gr. masc. n. *brotós*, a mortal human; Gr. fem. n. *kákkē*, dung; Gr. masc. n. *enoikos*, inhabitant; N.L. masc. n. *Brotocaccenecus*, a microbe from the faeces of humans).

The genus shared the highest 16S rRNA gene sequence similarities to species within family *Oscillospiraceae* (*Oscillibacter valericigenes*, 94.36%; *Oscillibacter ruminantium*, 94.29%). GTDB-Tk classified this isolate to the genus 'CAG-83' within family *Oscillospiraceae*. The POCP value below 50% to both *O. valericigenes* (the type species of this genus) and *O. ruminantium* (both were also the closest relatives based on the genome tree) confirmed the novel genus status. The type species is *Brotocaccenecus cirricatena*.

Description of *Brotocaccenecus cirricatena* sp. nov. *Brotocaccenecus cirricatena* (cir.ri.ca.te'na. L. masc. n. *cirrus*, a curl; L. fem. n. *catena*, a chain; N.L. fem. n. *cirricatena*, the curly chain, pertaining to the cell shape).

The species has all features of the genus. Cells are long curvy rods (>2 μm length) that grow under anaerobic conditions on YCFA agar medium. Genome analysis predicted the ability to utilize starch and produce acetate, butyrate, propionate, L-glutamate and riboflavin (vitamin B2). Antibiotic resistance genes related to tetracycline-resistant ribosomal protection protein (ARO:0000002) were detected. The G + C content of genomic DNA is 57.5 mol.%. The type strain, CLA-AA-H272^T, was most prevalent in pig gut microbiota (76.2% of 1000 samples positive), followed by wastewater (59.7%) and human gut microbiota (59.4%). It was isolated from the faeces of a healthy 28-year-old man.

Description of *Hominiventricola* gen. nov. *Hominiventricola* (Ho.mi.ni.ven.tri'co.la. L. masc. n. *homo*, a human being; L. masc. n. *venter*, the belly; N.L. masc./fem. suffix *-cola*, an inhabitant of; N.L. masc. n. *Hominiventricola*, a microbe from the intestine of humans).

The highest 16S rRNA gene sequence similarities to the isolate were species within family *Lachnospiraceae* (92.24%–93.88%), such as *Clostridium herbivorans* (still assigned to family *Clostridiaceae* in LPSN) and *Coprococcus comes*. GTDB-Tk classified the isolate to the genus 'UBA7182' within family *Lachnospiraceae*. Since no genome exists for *C. herbivorans*, genome comparison could not be performed. The POCP value of 44.25% to *Fusicatenibacter saccharivorans* (the type species of this genus and the closest relative based on the genome tree) and 25.0% to *C. butyricum* (type species of the genus *Clostridium*) confirmed the novel genus status within the *Lachnospiraceae*. The type species is *Hominiventricola filiformis*.

Description of *Hominiventricola filiformis* sp. nov. *Hominiventricola filiformis* (fi.li.for'mis. L. neut. n. *filum*, a string; L. fem. n. *forma*, a shape; N.L. masc. adj. *filiformis*, string-shaped).

The species has all features of the genus. It grows as thin long rods, almost string-like in modified Gifu Anaerobic Medium under anaerobic conditions. Genome analysis predicted the ability to utilize glucose, arbutin, salicin, trehalose, starch, sulfide and L-serine. The genes for production of acetate, propionate, L-cysteine and L-glutamate were also identified, along with the genes encoding for cobalamin (vitamin B12) and folate biosynthesis. No antibiotic resistance genes were detected. The G + C content is 45.1 mol.%. The type strain, CLA-AA-H276^T (=DSM 113464^T), was most prevalent in human gut microbiota (48.5% of 1000 samples positive), followed by wastewater (29.3%) and chicken gut microbiota (27.9%). It was isolated from the faeces of a healthy 30-year-old man.

Description of *Hominifimenecus* gen. nov. *Hominifimenecus* (Ho.mi.ni.fim.en.e'cus. L. masc. n. *homo*, a human being; L. neut. n. *fimum*, dung; Gr. masc. n. *enoikos*, inhabitant; N.L. masc. n. *Hominifimenecus*, a microbe from the faeces of humans).

The genus shared the highest 16S rRNA gene similarities to species within the family *Lachnospiraceae* (max. 93.35% sequence identity to *Lachnotalea glycerini*). GTDB-Tk classified the isolate as a novel genus within this family. This was confirmed by genome tree placement. Moreover, none of the close relatives shared a POCP value >50% to the isolate. The type species is *Hominifimenecus microfluidus*.

Description of *Hominifimenecus microfluidus* sp. nov. *Hominifimenecus microfluidus* (mi.cro.flu'i.dus. Gr. masc. adj. *mikros*, small; L. masc. adj. *fluidus*, flowing; N.L. masc. adj. *microfluidus*, pertaining to microfluidics, the method used for isolation of the species).

The species has all features of the genus. Cells are spindle-shaped coccobacilli, mostly 1.5–2.5 μm in length on BHI medium under anaerobic conditions. Genome analysis predicted the ability to utilize starch, sulfide and L-serine, while producing acetate, propionate, L-cysteine, L-glutamate and biosynthesise folate. Antibiotic resistance may be present due to detection of tetracycline-resistant ribosomal protection proteins (ARO:0000002). The molecular G + C content of genomic DNA is 49.1 mol.%. The type strain, CLA-AA-H215^T (=DSM 112725^T), was most prevalent in wastewater (26.6% of 1000 samples positive), followed by pig gut microbiota (22.2%) and human gut microbiota (15.4%). It was isolated from the faeces of a healthy 26-year-old woman.

Description of *Hominisplanchenecus* gen. nov. *Hominisplanchenecus* (Ho.mi.ni.splanch.en.e'cus. L. masc. n. *homo*, a human being; Gr. neut. n. *splánchnon*, guts; Gr. masc. n. *enoikos*, inhabitant; N.L. masc. n. *Hominisplanchenecus*, a microbe from the intestines of humans).

The closest taxonomical neighbours are species within the family *Lachnospiraceae*, with max. 95.25% 16S rRNA gene sequence identity to *Ruminococcus lactaris* (still assigned to family *Oscillospiraceae* in LPSN). GTDB-Tk assigned the genome to the genus 'CAG-56' within the family *Lachnospiraceae*. The closest relatives based on genome tree analysis were *Murimonas intestini* (type species) and *Robinsoniella peoriensis* (type species). None of these species, as well as *R. lactaris* had POCP values >50% to the isolate, confirming the novel genus status within family *Lachnospiraceae*. The type species is *Hominisplanchenecus faecis*.

Description of *Hominisplanchenecus faecis* sp. nov. *Hominisplanchenecus faecis* (fae'cis. L. gen. n. *faecis*, of faeces, from which the organism was isolated).

The species has all features of the genus. Cells are small rods with slightly pointy ends that tend to form pairs (length 1.0–2.0 μm) on BHI medium under anaerobic conditions. Genome analysis predicted the ability to utilize glucose and starch. The genes for production of acetate, propionate, L-glutamate and biosynthesis of cobalamin (vitamin B12), folate and riboflavin (vitamin B2) were identified. No antibiotic resistance genes were detected. The G + C content is 44.1 mol.%. The type strain, CLA-AA-H246^T (=DSM 113194^T), was most prevalent in chicken gut microbiota (44.3% of 1000 samples positive), followed by wastewater (39.7%) and human gut microbiota (34.1%). It was isolated from the faeces of a healthy 36-year-old woman.

Acknowledgements

We are grateful to Ntana Kousetzi (RWTH University Hospital) for outstanding technical assistance with sequencing, to Marzena Wyschkon (Leibniz Institute DSMZ) for the processing of bacterial strains, and to Aharon Oren (The Hebrew University of Jerusalem) for proof-reading the Latin bacterial names.

Funding Information

Deutsche Forschungsgemeinschaft (DFG): Project no. 403224013 (SFB1382) to TC; Project no. 460129525 (NFDI4Microbiota) to TC and JO

RWTH Aachen University (Exploratory Research Space program) OPSF442 to DK and TC

Authors Contributions

A.A. and T.C. conceived the studies. A.A., A.V., N.T. and E.M.B. performed experimental work. A.A., N.T. and T.C.A.H. performed sequence and bioinformatics analyses. A.A., T.C.A.H., A.V. and T.C. interpreted the data.

A.A., T.C.A.H., N.T., T.R. and B.A. curated data. J.O. provided access to essential materials and infrastructures. J.O., D.K. and T.C. secured funding. T.C. supervised the project. A.A. and T.C. generated figures and wrote the manuscript. T.C.A.H. critically reviewed the manuscript. All authors reviewed the final manuscript and agreed with its content.

Data Availability Statement

The 16S rRNA gene amplicon data generated in the present study are accessible in NCBI via project PRJNA767532. Accession numbers for the 16S rRNA gene sequences and draft genomes from isolates are listed in Supplementary Table S2 available via https://github.com/ClavelLab/SCD.

References

Almeida, A., Mitchell, A.L., Boland, M., Forster, S.C., Gloor, G.B., Tarkowska, A., *et al*. (2019) A new genomic blueprint of the human gut microbiota. *Nature* **568**: 499–504. https://doi.org/10.1038/s41586-019-0965-1.

Almeida, A., Nayfach, S., Boland, M., Strozzi, F., Beracochea, M., Shi, Z.J., *et al*. (2021) A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat Biotechnol* **39**: 105–114. https://doi.org/10.1038/s41587-020-0603-3.

Attebery, H.R., and Finegold, S.M. (1969) Combined screw-cap and rubber-stopper closure for Hungate tubes (pre-reduced anaerobically sterilized roll tubes and liquid media). *Appl Microbiol* **18**: 558–561. https://doi.org/10.1128/am.18.4.558-561.1969.

Bahram, M., Hildebrand, F., Forslund, S.K., Anderson, J.L., Soudzilovskaia, N.A., Bodegom, P.M., *et al*. (2018) Structure and function of the global topsoil microbiome. *Nature* **560**: 233–237. https://doi.org/10.1038/s41586-018-0386-6.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., *et al*. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**: 455–477. https://doi.org/10.1089/cmb.2012.0021.

Bello, M.G.D., Knight, R., Gilbert, J.A., and Blaser, M.J. (2018) Preserving microbial diversity. *Science* **362**: 33–34. https://doi.org/10.1126/science.aau8816.

Beresford-Jones, B.S., Forster, S.C., Stares, M.D., Notley, G., Viciani, E., Browne, H.P., *et al*. (2022) The mouse gastrointestinal bacteria catalogue enables translation between the mouse and human gut microbiotas via functional mapping. *Cell Host Microbe* **30**: 124–138.e8. https://doi.org/10.1016/j.chom.2021.12.003.

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.-C.C., Charles, T., *et al*. (2020) Microbiome definition re-visited: old concepts and new challenges. *Microbiome* **8**: 103. https://doi.org/10.1186/s40168-020-00875-0.

Berry, D., Ben, M.K., Wagner, M., and Loy, A. (2011) Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. *Appl Environ Microbiol* **77**: 7846–7849. https://doi.org/10.1128/AEM.05220-11.

Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Browne, H.P., Forster, S.C., Anonye, B.O., Kumar, N., Neville, B.A., Stares, M.D., *et al*. (2016) Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature* **533**: 543–546. https://doi.org/10.1038/nature17645.

Bushnell, B. (2014) BBMap: A Fast, Accurate, Splice-Aware Aligner.

Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P., and Parks, D. H. (2019) GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**: 1925–1927. https://doi.org/10.1093/bioinformatics/btz848.

Clavel, T., Charrier, C., and Haller, D. (2013) Streptococcus danieliae sp. nov., a novel bacterium isolated from the caecum of a mouse. *Arch Microbiol* **195**: 43–49. https://doi.org/10.1007/s00203-012-0846-6.

Clavel, T., Lagkouvardos, I., and Hiergeist, A. (2016) Microbiome sequencing: challenges and opportunities for molecular medicine. *Expert Rev Mol Diagn* **16**: 795–805. https://doi.org/10.1080/14737159.2016.1184574.

Durand, G.A., Cadoret, F., Lagier, J.C., Fournier, P.E., and Raoult, D. (2017a) Description of 'Gorbachella massiliensis' gen. nov., sp. nov., 'Fenollaria timonensis' sp. nov., 'Intestinimonas timonensis' sp. nov. and 'Collinsella ihuae' sp. nov. isolated from healthy fresh stools with culturomics. *New Microbes New Infect* **16**: 60–62. https://doi.org/10.1016/j.nmni.2017.01.005.

Durand, G.A., Pham, T., Ndongo, S., Traore, S.I., Dubourg, G., Lagier, J.-C.C., *et al*. (2017b) Blautia massiliensis sp. nov., isolated from a fresh human fecal sample and emended description of the genus Blautia. *Anaerobe* **43**: 47–55. https://doi.org/10.1016/j.anaerobe.2016.12.001.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., *et al*. (2005) Microbiology: diversity of the human intestinal microbial flora. *Science* **308**: 1635–1638. https://doi.org/10.1126/science.1110591.

Edgar, R.C. (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* **10**: 996–998. https://doi.org/10.1038/nmeth.2604.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**: 2194–2200. https://doi.org/10.1093/bioinformatics/btr381.

Faith, J.J., Guruge, J.L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A.L., *et al*. (2013) The long-term stability of the human gut microbiota. *Science* **341**: 1237439. https://doi.org/10.1126/science.1237439.

Floch, M.H., Gershengoren, W., and Freedman, L.R. (1968) Methods for the quantitative study of the aerobic and anaerobic intestinal bacterial flora of man. *Yale J Biol Med* **41**: 50–61.

Fodor, A.A., DeSantis, T.Z., Wylie, K.M., Badger, J.H., Ye, Y., Hepburn, T., *et al*. (2012) The "Most Wanted" taxa from the human microbiome for whole genome sequencing. *PLoS One* **7**: e41294. https://doi.org/10.1371/journal.pone.0041294.

Forster, S.C., Kumar, N., Anonye, B.O., Almeida, A., Viciani, E., Stares, M.D., *et al*. (2019) A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat Biotechnol* **37**: 186–192. https://doi.org/10.1038/s41587-018-0009-7.

Godon, J.J., Zumstein, E., Dabert, P., Habouzit, F., and Moletta, R. (1997) Molecular microbial diversity of an anaerobic digestor as determined by small-subunit rDNA sequence analysis. *Appl Environ Microbiol* **63**: 2802–2813. https://doi.org/10.1128/aem.63.7.2802-2813.1997.

Goodman, A.L., Kallstrom, G., Faith, J.J., Reyes, A., Moore, A., Dantas, G., and Gordon, J.I. (2011) Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc Natl Acad Sci U S A* **108**: 6252–6257.

Greuter, D., Loy, A., Horn, M., and Rattei, T. (2016) ProbeBase-an online resource for rRNA-targeted oligonucleotide probes and primers: new features 2016. *Nucleic Acids Res* **44**: D586–D589. https://doi.org/10.1093/nar/gkv1232.

Groussin, M., Poyet, M., Sistiaga, A., Kearney, S.M., Moniz, K., Noel, M., *et al*. (2021) Elevated rates of horizontal gene transfer in the industrialized human microbiome. *Cell* **184**: 2053–2067.e18. https://doi.org/10.1016/j.cell.2021.02.052.

Haas, K.N., and Blanchard, J.L. (2020) Reclassification of the Clostridium clostridioforme and Clostridium sphenoides clades as Enterocloster gen. nov. and Lacrimispora gen. nov., including reclassification of 15 taxa. *Int J Syst Evol Microbiol* **70**: 23–34. https://doi.org/10.1099/ijsem.0.003698.

Hitch, T.C.A., Afrizal, A., Riedel, T., Kioukis, A., Haller, D., Lagkouvardos, I., *et al*. (2021a) Recent advances in culture-based gut microbiome research. *Int J Med Microbiol* **311**: 151485. https://doi.org/10.1016/j.ijmm.2021.151485.

Hitch, T.C.A., Riedel, T., Oren, A., Overmann, J., Lawley, T. D., and Clavel, T. (2021b) Automated analysis of genomic sequences facilitates high-throughput and comprehensive description of bacteria. *ISME Commun* **1**: 16. https://doi.org/10.1038/s43705-021-00017-z.

Hu, B., Xu, B., Yun, J., Wang, J., Xie, B., Li, C., *et al*. (2020) High-throughput single-cell cultivation reveals the underexplored rare biosphere in deep-sea sediments along the Southwest Indian Ridge. *Lab Chip* **20**: 363–372. https://doi.org/10.1039/C9LC00761J.

Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., and Aluru, S. (2018) High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* **9**: 5114. https://doi.org/10.1038/s41467-018-07641-9.

Janney, A., Powrie, F., and Mann, E.H. (2020) Host–microbiota maladaptation in colorectal cancer. *Nature* **585**: 509–517. https://doi.org/10.1038/s41586-020-2729-3.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., and Glöckner, F.O. (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* **41**: e1. https://doi.org/10.1093/nar/gks808.

Knight, R., Vrbanac, A., Taylor, B.C., Aksenov, A., Callewaert, C., Debelius, J., *et al*. (2018) Best practices for analysing microbiomes. *Nat Rev Microbiol* **16**: 410–422. https://doi.org/10.1038/s41579-018-0029-9.

Kumar, S., Stecher, G., and Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* **33**: 1870–1874. https://doi.org/10.1093/molbev/msw054.

Lagier, J.C., Khelaifia, S., Alou, M.T., Ndongo, S., Dione, N., Hugon, P., *et al*. (2016) Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat Microbiol* **1**: 16203. https://doi.org/10.1038/nmicrobiol.2016.203.

Lagkouvardos, I., Fischer, S., Kumar, N., and Clavel, T. (2017a) Rhea: a transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons. *PeerJ* **5**: e2836. https://doi.org/10.7717/peerj.2836.

Lagkouvardos, I., Joseph, D., Kapfhammer, M., Giritli, S., Horn, M., Haller, D., and Clavel, T. (2016a) IMNGS: a comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. *Sci Rep* **6**: 33721. https://doi.org/10.1038/srep33721.

Lagkouvardos, I., Overmann, J., and Clavel, T. (2017b) Cultured microbes represent a substantial fraction of the human and mouse gut microbiota. *Gut Microbes* **8**: 493–503. https://doi.org/10.1080/19490976.2017.1320468.

Lagkouvardos, I., Pukall, R., Abt, B., Foesel, B.U., Meier-Kolthoff, J.P., Kumar, N., *et al*. (2016b) The Mouse Intestinal Bacterial Collection (miBC) provides host-specific insight into cultured diversity and functional potential of the gut microbiota. *Nat Microbiol* **1**: 16131. https://doi.org/10.1038/nmicrobiol.2016.131.

Letunic, I., and Bork, P. (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* **49**: W293–W296. https://doi.org/10.1093/nar/gkab301.

Liu, C., Du, M.-X., Abuduaini, R., Yu, H.-Y., Li, D.-H., Wang, Y.-J., *et al*. (2021) Enlightening the taxonomy darkness of human gut microbiomes with a cultured biobank. *Microbiome* **9**: 119. https://doi.org/10.1186/s40168-021-01064-3.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**: D490–D495. https://doi.org/10.1093/nar/gkt1178.

Lynch, J.B., and Hsiao, E.Y. (2019) Microbiomes as sources of emergent host phenotypes. *Science* **365**: 1405–1409. https://doi.org/10.1126/science.aay0240.

McArthur, A.G., Waglechner, N., Nizam, F., Yan, A., Azad, M. A., Baylay, A.J., *et al*. (2013) The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother* **57**: 3348–3357. https://doi.org/10.1128/AAC.00419-13.

Meier-Kolthoff, J.P., Klenk, H.P., and Göker, M. (2014) Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* **64**: 352–356. https://doi.org/10.1099/ijs.0.056994-0.

Moore, W.E.C., and Holdeman, L.V. (1974) Human fecal Flora: the Normal Flora of 20 Japanese-Hawaiians. *Appl Microbiol* **27**: 961–979. https://doi.org/10.1128/am.27.5.961-979.1974.

Murray, A.E., Freudenstein, J., Gribaldo, S., Hatzenpichler, R., Hugenholtz, P., Kämpfer, P., *et al*. (2020) Roadmap for naming uncultivated Archaea and bacteria. *Nat Microbiol* **5**: 987–994. https://doi.org/10.1038/s41564-020-0733-x.

O'Keefe, S.J. (2016) Diet, microorganisms and their metabolites, and colon cancer. *Nat Rev Gastroenterol Hepatol* **13**: 691–706. https://doi.org/10.1038/nrgastro.2016.165.

Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. (1999) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **27**: 29–34. https://doi.org/10.1093/nar/27.1.29.

Ondov, B.D., Treangen, T.J., Melsted, P., Mallonee, A.B., Bergman, N.H., Koren, S., and Phillippy, A.M. (2016) Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol* **17**: 132. https://doi.org/10.1186/s13059-016-0997-x.

Overmann, J. (2013) Principles of enrichment, isolation, cultivation, and preservation of prokaryotes. In *The Prokaryotes*, Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H., and Stackebrandt, E. (eds). Berlin, Heidelberg: Springer, pp. 149–207. https://doi.org/10.1007/978-3-642-30194-0_7.

Overmann, J., Abt, B., and Sikorski, J. (2017) Present and future of culturing bacteria. *Annu Rev Microbiol* **71**: 711–730. https://doi.org/10.1146/annurev-micro-090816-093449.

Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.A., and Hugenholtz, P. (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* **36**: 996–1004. https://doi.org/10.1038/nbt.4229.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043–1055. https://doi.org/10.1101/gr.186072.114.

Pruesse, E., Peplies, J., Glöckner, F.O., and Glockner, F.O. (2012) SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**: 1823–1829. https://doi.org/10.1093/bioinformatics/bts252.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., and Glockner, F.O. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**: 7188–7196. https://doi.org/10.1093/nar/gkm864.

Qin, Q.-L., Xie, B.-B., Zhang, X.-Y., Chen, X.-L., Zhou, B.-C., Zhou, J., *et al*. (2014) A proposed genus boundary for the prokaryotes based on genomic insights. *J Bacteriol* **196**: 2210–2215. https://doi.org/10.1128/JB.01688-14.

Reitmeier, S., Hitch, T.C.A., Treichel, N., Fikas, N., Hausmann, B., Ramer-Tait, A.E., *et al*. (2021) Handling of spurious sequences affects the outcome of high-throughput 16S rRNA gene amplicon profiling. *ISME Commun* **1**: 31. https://doi.org/10.1038/S43705-021-00033-Z.

Riba, J., Gleichmann, T., Zimmermann, S., Zengerle, R., and Koltay, P. (2016) Label-free isolation and deposition of single bacterial cells from heterogeneous samples for clonal culturing. *Sci Rep* **6**: 32837. https://doi.org/10.1038/srep32837.

Rutgeerts, P., Peeters, M., Hiele, M., Vantrappen, G., Pennincx, F., Aerts, R., *et al*. (1991) Effect of faecal stream diversion on recurrence of Crohn's disease in the neoterminal ileum. *Lancet* **338**: 771–774. https://doi.org/10.1016/0140-6736(91)90663-A.

Salazar, G., Paoli, L., Alberti, A., Huerta-Cepas, J., Ruscheweyh, H.-J., Cuenca, M., *et al*. (2019) Gene expression changes and community turnover differentially shape the Global Ocean Metatranscriptome. *Cell* **179**: 1068–1083.e21. https://doi.org/10.1016/j.cell.2019.10.014.

Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**: 2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

Sender, R., Fuchs, S., and Milo, R. (2016) Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol* **14**: e1002533. https://doi.org/10.1371/journal.pbio.1002533.

Sepich-Poore, G.D., Zitvogel, L., Straussman, R., Hasty, J., Wargo, J.A., and Knight, R. (2021) The microbiome and human cancer. *Science* **371**: eabc4552. https://doi.org/10.1126/science.abc4552.

Seshadri, R., Leahy, S.C., Attwood, G.T., Teh, K.H., Lambie, S.C., Cookson, A.L., *et al*. (2018) Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. *Nat Biotechnol* **36**: 359–367. https://doi.org/10.1038/nbt.4110.

Sonnenburg, E.D., Smits, S.A., Tikhonov, M., Higginbottom, S.K., Wingreen, N.S., and Sonnenburg, J.L. (2016) Diet-induced extinctions in the gut microbiota compound over generations. *Nature* **529**: 212–215. https://doi.org/10.1038/nature16504.

Sonnenburg, J.L., and Sonnenburg, E.D. (2019) Vulnerability of the industrialized microbiota. *Science* **366**: eaaw9255. https://doi.org/10.1126/SCIENCE.AAW9255.

Stephen, A.M., and Cummings, J.H. (1980) The microbial contribution to human faecal mass. *J Med Microbiol* **13**: 45–56. https://doi.org/10.1099/00222615-13-1-45.

Thomas, A.M., and Segata, N. (2019) Multiple levels of the unknown in microbiome research. *BMC Biol* **17**: 48. https://doi.org/10.1186/s12915-019-0667-z.

van Nood, E., Vrieze, A., Nieuwdorp, M., Fuentes, S., Zoetendal, E.G., de Vos, W.M., *et al*. (2013) Duodenal infusion of donor feces for recurrent Clostridium difficile. *N Engl J Med* **368**: 407–415. https://doi.org/10.1056/NEJMoa1205037.

Visconti, A., Le Roy, C.I., Rosa, F., Rossi, N., Martin, T.C., Mohney, R.P., *et al*. (2019) Interplay between the human gut microbiome and host metabolism. *Nat Commun* **10**: 1–10. https://doi.org/10.1038/s41467-019-12476-z.

Watterson, W.J., Tanyeri, M., Watson, A.R., Cham, C.M., Shan, Y., Chang, E.B., *et al*. (2020) Droplet-based high-throughput cultivation for accurate screening of antibiotic resistant gut microbes. *Elife* **9**: 1–22. https://doi.org/10.7554/eLife.56998.

Wiegand, S., Dam, H.T., Riba, J., Vollmers, J., and Kaster, A.K. (2021) Printing microbial dark matter: using single cell dispensing and genomics to investigate the Patescibacteria/candidate phyla radiation. *Front Microbiol* **12**: 1512. https://doi.org/10.3389/fmicb.2021.635506.

Wylensek, D., Hitch, T.C.A., Riedel, T., Afrizal, A., Kumar, N., Wortmann, E., *et al*. (2020) A collection of

bacterial isolates from the pig intestine reveals functional and taxonomic diversity. *Nat Commun* **11**: 1–26. https://doi.org/10.1038/s41467-020-19929-w.

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.-H.H., *et al*. (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* **12**: 635–645. https://doi.org/10.1038/nrmicro3330.

Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., *et al*. (2014) The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res* **42**: D643–D648. https://doi.org/10.1093/nar/gkt1209.

Yoon, S.H., Ha, S., Lim, J., Kwon, S., and Chun, J. (2017a) A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* **110**: 1281–1286. https://doi.org/10.1007/s10482-017-0844-4.

Yoon, S.H., Ha, S.M., Kwon, S., Lim, J., Kim, Y., Seo, H., and Chun, J. (2017b) Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* **67**: 1613–1617. https://doi.org/10.1099/ijsem.0.001755.

Zehavi, T., Probst, M., and Mizrahi, I. (2018) Insights into culturomics of the rumen microbiome. *Front Microbiol* **9**: 1–10. https://doi.org/10.3389/fmicb.2018.01999.

Zenner, C., Hitch, T.C.A., Riedel, T., Wortmann, E., Tiede, S., Buhl, E.M., *et al*. (2021) Early-life immune system maturation in chickens using a synthetic Community of Cultured gut Bacteria. *mSystems* **6**: e01300–e01320. https://doi.org/10.1128/mSystems.01300-20.

Zhou, N., Sun, Y., Chen, D., Du, W., Yang, H., and Liu, S. (2019) Harnessing microfluidic streak plate technique to investigate the gut microbiome of Reticulitermes chinensis. *Microbiology* **8**: e00654. https://doi.org/10.1002/mbo3.654.

Zou, Y., Xue, W., Luo, G., Deng, Z., Qin, P., Guo, R., *et al*. (2019) 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nat Biotechnol* **37**: 179–185. https://doi.org/10.1038/s41587-018-0008-8.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Supplementary Table S1. 16S rRNA gene amplicon sequencing of Experiment Type 1.

Supplementary Table S2. List of bacteria isolated from experiment type 1.

Supplementary Table S3. 16S rRNA gene amplicon sequencing results of SCD Experiment Type 2.

Fig. S1. Phylogenomic tree showing the placement of strain CLA-AA-H232T within members of the familiy *Oscillopiraceae*. The tree was created as detailed in Protologger (Hitch *et al*., 2021b).