

Alkaliphilus flagellatus sp. nov., *Butyricicoccus intestinisimiae* sp. nov., *Clostridium mobile* sp. nov., *Clostridium simiarum* sp. nov., *Dysosmobacter acutus* sp. nov., *Paenibacillus brevis* sp. nov., *Peptoniphilus ovalis* sp. nov. and *Tissierella simiarum* sp. nov., isolated from monkey faeces

Dan-Hua Li¹, Rexiding Abuduaini^{1,2}, Meng-Xuan Du¹, Yu-Jing Wang^{1,2}, Hong-He Chen¹, Nan Zhou¹, Hai-Zhen Zhu¹, Yong Lu³, Pei-Jun Yu^{2,3}, Yun-Peng Yang³, Cheng-Ying Jiang^{1,2,3}, Qiang Sun³, Chang Liu^{1,*} and Shuang-Jiang Liu^{1,2,4,*}

Abstract

Non-human primates harbour diverse microbiomes in their guts. As a part of the China Microbiome Initiatives, we cultivated and characterized the gut microbiome of cynomolgus monkeys (*Macaca fascicularis*). In this report, we communicate the characterization and taxonomy of eight bacterial strains that were obtained from faecal samples of captive cynomolgus monkeys. The results revealed that they represented eight novel bacterial species. The proposed names of the eight novel species are *Alkaliphilus flagellatus* (type strain MSJ-5^T=CGMCC 1.45007^T=KCTC 15974^T), *Butyricicoccus intestinisimiae* MSJd-7^T (MSJd-7^T=CGMCC 1.45013^T=KCTC 25112^T), *Clostridium mobile* (MSJ-11^T=CGMCC 1.45009^T=KCTC 25065^T), *Clostridium simiarum* (MSJ-4^T=CGMCC 1.45006^T=KCTC 15975^T), *Dysosmobacter acutus* (MSJ-2^T=CGMCC 1.32896^T=KCTC 15976^T), *Paenibacillus brevis* MSJ-6^T (MSJ-6^T=CGMCC 1.45008^T=KCTC 15973^T), *Peptoniphilus ovalis* (MSJ-1^T=CGMCC 1.31770^T=KCTC 15977^T) and *Tissierella simiarum* (MSJ-40^T=CGMCC 1.45012^T=KCTC 25071^T).

INTRODUCTION

Gastrointestinal tracts accommodate diverse microbes, and those microbes together in a host gastrointestinal tract are called gut microbiomes (GMs) [1]. Many efforts have been made to characterize the microbial diversities of human [2–6] and animal GMs [7, 8], by culture-dependent and/or -independent methods [9–11]. Non-human primates (NHPs) are the most biologically relevant animal models for human studies [12]. The compositions and dynamics of NHP GMs were evaluated, and members of the genera *Bacteroides*, *Bifidobacterium*, *Eubacterium*, *Fusobacterium*, *Lactobacillus* and *Streptococcus* were cultivated and reported from the gastrointestinal tract of NHPs [13–17]. Based on analyses of major and large-scale investigations of human GMs [2–6], there are 5000–6000 bacterial species associated with humans. An exploration of cultivated human gut bacterial species diversity revealed that about 1500 bacterial species are recorded with valid and correct names, and that more gut bacterial species have been

Author affiliations: ¹State Key Laboratory of Microbial Resources and Environmental Microbiology Research Center at Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, PR China; ²University of Chinese Academy of Sciences, Beijing 100049, PR China; ³Institute of Neuroscience, CAS Center for Excellence in Brain Science and Intelligence Technology, State Key Laboratory of Neuroscience, CAS Key Laboratory of Primate Neurobiology, Chinese Academy of Sciences, 200031, Shanghai, PR China; ⁴State Key Laboratory of Microbial Biotechnology, Shandong University, Qingdao 266237, PR China.

***Correspondence:** Shuang-Jiang Liu, liusj@im.ac.cn; liusj@sdu.edu.cn; Chang Liu, liuc@im.ac.cn

Keywords: non-human primates; gut microbiome; culturomics; polyphasic taxonomy; cynomolgus monkeys (*Macaca fascicularis*).

Abbreviations: ANI, average nucleotide identity; dDDH, digital DNA–DNA hybridization; GM, gut microbiome; NHP, non-human primate; SCFA, short-chain fatty acid.

The NCBI/NMDC accession numbers of the 16S rRNA gene and genome sequences for the eight type strains are: (i) MSJ-1^T, 16S rRNA sequence: MZ310594 and NMDCN0000NQV; genome sequence: JAHLQ000000000/NMDC60018343; (ii) MSJ-2^T, 16S rRNA sequence: MZ310595/NMDCN0000NR0; genome sequence: JAHLQ000000000/NMDC60018344; (iii) MSJ-4^T, 16S rRNA sequence: MZ310597/NMDCN0000NR2; genome sequence: JAHLQ000000000/NMDC60018346; (iv) MSJ-5^T, 16S rRNA sequence: MZ310598/NMDCN0000NR3; genome sequence: JAHLQ000000000/NMDC60018347; (v) MSJ-6^T, 16S rRNA sequence: MZ310599/NMDCN0000NR4; genome sequence: JAHLQJ000000000/NMDC60018348; (vi) MSJd-7^T, 16S rRNA sequence: MZ310600/NMDCN0000NR5; genome sequence: JAHLQI000000000/NMDC60018349; (vii) MSJ-11^T, 16S rRNA sequence: MZ310603/NMDCN0000NR8; genome sequence: JAHLQF000000000/NMDC60018352; (viii) MSJ-40^T, 16S rRNA sequence: MZ310625/NMDCN0000NRU; genome sequence: JAHLP000000000/NMDC60018374.

Three supplementary figures are available with the online version of this article.

005276 © 2022 The Authors



cultivated but not characterized or nominated [2]. Those cultivated but unnamed bacterial species have remained as ‘uncultivated’ in databases such as SILVA [18] and the Genome Taxonomy Database [19], and they are repeatedly claimed as ‘novel bacteria’ in later studies. Bacterial cultivation, characterization with polyphasic methods, and deposits in culture collections are essential to nominate a bacterial name following the rules of International Code of Nomenclature of Prokaryotes. Considering the large numbers (usually more than thousands) of bacterial isolates from one microbiome study, it is a challenge to characterize and nominate all bacterial isolates from microbiome studies.

Members of the family *Clostridiaceae* [20], such as *Clostridium* and *Alkaliphilus*, are frequently detected in guts and they play important roles in host health. *Clostridium* species, a group of 153 validly published and correct specific names (<https://lpsn.dsmz.de/genus/clostridium>) and represented by the type species *Clostridium butyricum* [21], are Gram-positive, obligately anaerobic rods and form oval or spherical endospores. The genus *Alkaliphilus* is also a member of family *Clostridiaceae* [22] and contains seven validly published and correct specific names (<https://lpsn.dsmz.de/genus/alkaliphilus>). *Alkaliphilus* species are strictly anaerobic and their cells are usually straight to slightly curved rods. The genus *Paenibacillus* of the family *Paenibacillaceae* [23] is one of the largest genera of prokaryotes and members are widely distributed in natural environments and animal and human GMs. At the time of writing, 270 species names are validly published (<https://lpsn.dsmz.de/genus/paenibacillus>). *Paenibacillus* species are usually straight to slightly curved rods and motile with flagella. The type species is *Paenibacillus polymyxa* and was isolated from decomposing plant materials and humus-rich soils. The genus *Peptoniphilus* is a member of the family *Peptoniphilaceae* [24]. *Peptoniphilus* species are non-spore-forming, Gram-positive, obligately anaerobic cocci, and the type species is *Peptoniphilus asaccharolyticus*. At the time of writing, 20 species of the genus *Peptoniphilus* have been validly published (<https://lpsn.dsmz.de/genus/peptoniphilus>). The genera *Dysosmobacter* and *Butyricoccus* are members of the family *Oscillospiraceae* and were proposed by Le Roy *et al.* [25] and Eeckhaut *et al.* [26], respectively. So far, only the type species *Dysosmobacter welbionis* is described (<https://lpsn.dsmz.de/genus/dysosmobacter>). *D. welbionis* is an obligately anaerobic, non-spore-forming and non-motile rod. The genus *Butyricoccus* has four validly and correctly named species (<https://lpsn.dsmz.de/genus/butyricoccus>), and they were isolated from human or animal faeces. Cells of the genus *Butyricoccus* are anaerobic, non-motile and coccoid; the type species is *Butyricoccus pullicaecorum*. The genus *Tissierella* belongs to the family *Tissierellaceae* [27]. *Tissierella* species are obligately anaerobic, Gram-negative, non-spore forming rods. At the time of writing, five specific names have been validly published (<https://lpsn.dsmz.de/genus/tissierella>), and the type species is *Tissierella praeacuta*.

The China Microbiome Initiatives (CMI) integrated multiple studies of human and animal GMs and environmental microbiomes [28]. As a part of the CMI, we cultivated and characterized the GM of cynomolgus monkeys (*Macaca fascicularis*). In this report, we present the characterization and taxonomy of eight bacterial strains that were obtained from faecal samples of captive cynomolgus monkeys. The eight bacterial strains are affiliated with five families and were identified as new members of the genera *Peptoniphilus*, *Dysosmobacter*, *Clostridium*, *Alkaliphilus*, *Paenibacillus*, *Butyricoccus* and *Tissierella*.

METHODS

Sample collection and treatment

All faecal samples were from cynomolgus monkeys (*M. fascicularis*) at the experimental animal centre of the Institute of Neuroscience, Chinese Academy of Sciences, Suzhou, PR China. Fresh faecal samples were collected and maintained in airtight bags on dry ice, and were delivered immediately to the laboratory. The samples were diluted with sterile PBS and filtered through a 40 µm cell strainer and were treated with 70% ethanol or heated at 85 °C for 30 min, as described in the literatures [29, 30].

Culture media, bacterial isolation and cultivation

The following media were used for bacterial cultivation: FAB (Fastidious anaerobe broth) medium (LA4550, Solarbio) and YCFA (yeast extract, casein hydrolysate, fatty acids) medium [31], modified GAM (Gifu anaerobic medium) medium (mGAM) [2, 32], and modified R-medium named by Dione N *et al* [33]. The mGAM medium (per 1 l) contained 10 g casitone, 3 g soya peptone, 15 g proteose peptone, 13.5 g digested serum, 5 g yeast extract, 2 g beef extract powder, 1.2 g liver extract, 0.3 g soluble starch, 0.5 g L-cysteine, 0.5 g L-arginine, 0.3 g L-tryptophan, 2 g NaHCO₃, 2.5 g KH₂PO₄, 3 g NaCl, 0.15 g CH₂(SH)COONa, 2.46 g CH₃COONa, 0.01 g haemin, 0.001 g resazurin, 0.3 g glucose, 0.3 g D-galactose, 0.3 g cellobiose, 0.3 g mannose, 0.3 g fructose, 0.3 g rhamnose, 0.3 g palatinose, 0.3 g inulin, 15 g agar, adjusted pH to 7.2, sterilized at 115 °C for 25 min. The modified R medium was prepared from two solutions that were prepared, sterilized and separated: solution A (per 900 ml) consisted of 6 g casein hydrolysate, 5 g peptone, 5 g yeast extract, 1 g glucose, 1 g inulin, 1 g D-fructose, 1 g cellobiose, 1.5 g NaCl, 0.1 g MgSO₄ · H₂O, 5 ml haemin (0.1%, w/v), 1 ml resazurin (0.1%, w/v), 20 ml (2%, v/v) rumen fluid, 15 g agar, adjusted pH to 7.2, sterilized at 112 °C for 15 min. Solution B (per 100 ml) consisted of 0.4 g L-cysteine, 1 g ascorbic acid, 0.1 g glutathione, 2 g α-ketoglutarate, 0.45 g K₂HPO₄, 0.9 g KH₂PO₄, adjusted pH to 7.2, filtered using a 0.2 µm micro filter.

Sterilized agar plates were inoculated with dilutions of pretreated faecal samples and incubated at 37 °C under strictly anaerobic conditions with N₂ (85 %), H₂ (10 %) and CO₂ (5 %) in an anaerobic chamber (Electrotek AW400SG workstation). Colonies

appeared after cultivation for 2, 5 and 10 days were picked and re-streaked on agar plates of same media. Bacterial purity was evaluated by observation of morphology, 16S rRNA gene and genome sequencing.

Cell morphology observation and chemotaxonomic determinations

Cell morphology was determined by transmission electron microscopy (JEM-1400, JEOL). The utilization of carbon sources was determined using the 96-well Biolog AN MicroPlate that contained 95 different carbon substrates [34]. Bacteria strains were cultured in liquid mGAM medium for 2 days, then cells were harvested. Cellular fatty acids were extracted and methylated according to the standard MIDI protocol (Sherlock Microbial Identification System, version 6.0). The identification was performed by GC (HP 6890 Series GC System; Agilent) [35]. Polar lipids were separated by two-dimensional thin-layer chromatography (TLC) plates coated with silica gel, 1010 cm; Merck). Chromatography was performed using chloroform–methanol–water (65:25:4, by vol.) for the first dimension, followed by chloroform–methanol–acetic acid–water (80:12:15:4, by vol.) for the second dimension [36]. Total lipids were detected with 10% ethanolic molybdatophosphoric acid (Sigma). Aminolipids were detected with 0.4% solution of ninhydrin (Sigma) in butanol. Phospholipids were detected with Zinzadze reagent (molybdenum blue spray reagent, 1.3%; Sigma) and glycolipids were detected with 0.5% α -naphthol sulphuric acid reagent.

Fermentative production of short-chain fatty acids

Bacterial strains were cultivated for 72 h in mGAM broth at 37 °C under strictly anaerobic conditions. Short-chain fatty acids (SCFAs) were measured using GC-MS. Culture (1 ml) was extracted with 1 ml ethyl acetate. The supernatant liquid was prepared for GC-MS analysis, which was performed on a GCMS-QP2010 Ultra with an auto sampler (Shimadzu) and the DB-wax capillary column (30 m, 0.25 mm i.d., 0.25 μ m film thickness, Agilent Technologies). The temperature of oven was programmed from 35 to 130 °C at 5 °C min⁻¹ gradient, to 230 °C at 30 °C min⁻¹ gradient, with 16 min hold. Injection of 2 μ l sample was performed at 230 °C. The carrier gas, helium, flowed at 1.0 ml min⁻¹. Electronic impact was recorded at 70 eV.

16S rRNA gene sequencing and phylogenetic analysis

Complete 16S rRNA gene sequences of isolates were obtained using the universal primers 27F (5'-AGAGTTTGATCCTG-GCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3') [37]. 16S rRNA gene sequences similarities were determined using the EzBioCloud server [38]. Multiple alignments of sequences were performed using the Clustal W [39]. The phylogenetic trees were reconstructed by the neighbour-joining method [40] according to Kimura's two-parameter model [41] in MEGA X [42], by the maximum-likelihood method [43] based on the Tamura–Nei model, and the maximum-parsimony method [44] based on the subtree-pruning-regrafting search method. The statistical reliability of the trees was calculated by bootstrap analysis with 1000 replications [45].

Genome sequencing and analysis

Genomic DNA was extracted using the Wizard Genomic DNA Purification Kit (Promega) and the genomic DNA library was sequenced on an Illumina HiSeq X-ten platform. All good-quality paired reads were assembled using SPAdes software (version 3.9.0) [46]. The average nucleotide identity (ANI) values with closely related and available genomes were calculated using OAT software at www.ezbiocloud.net/sw/oat along with the UPGMA dendrogram (unweighted pair group method with arithmetic mean) [47]. The genomic distances, digital DNA–DNA hybridization (dDDH), were calculated by using the Genome-to-Genome Distance Calculator (<http://ggdc.dsmz.de/>) [48]. Genome analysis using the Check M indicated that the genomes of eight strains were not contaminated [49].

Culture preservation

Bacterial strains were cultured in liquid medium for 2 days. We stored our own cultures (1 ml) in the lab by addition of an equal volume of 65% (v/v) glycerol (1 ml), and was storing at –80 °C for long-term preservation. All type strains assigned by this study were deposited at China General Microbiological Culture Collection Center (CGMCC) and the Korean Collection for Type Culture (KCTC), and strain numbers are included in the species descriptions.

RESULTS AND DISCUSSION

Source and isolation of the bacteria

Strains MSJ-1^T, MSJ-2^T, MSJ-4^T, MSJ-5^T, MSJ-6^T, MSJd-7^T, MSJ-11^T and MSJ-40^T were isolated from faeces samples of *Macaca fascicularis*. Strain MSJ-1^T was obtained from the sample after enrichment, strains MSJ-4^T and MSJ-5^T were obtained from samples after heat treatment (85 °C for 30 min), and strains MSJ-2^T, MSJ-6^T, MSJd-7^T, MSJ-11^T and MSJ-40^T were obtained from samples after 70% ethanol treatment for 30 min. Strains MSJ-1^T, MSJ-2^T, MSJ-4^T and MSJ-11^T were successfully cultivated first with FAB, strain MSJ-5^T first from YCEFA, strain MSJd-7^T first from modified mGAM, and MSJ-40^T first from modified R media. However, we later demonstrated that they all grew with mGAM medium.

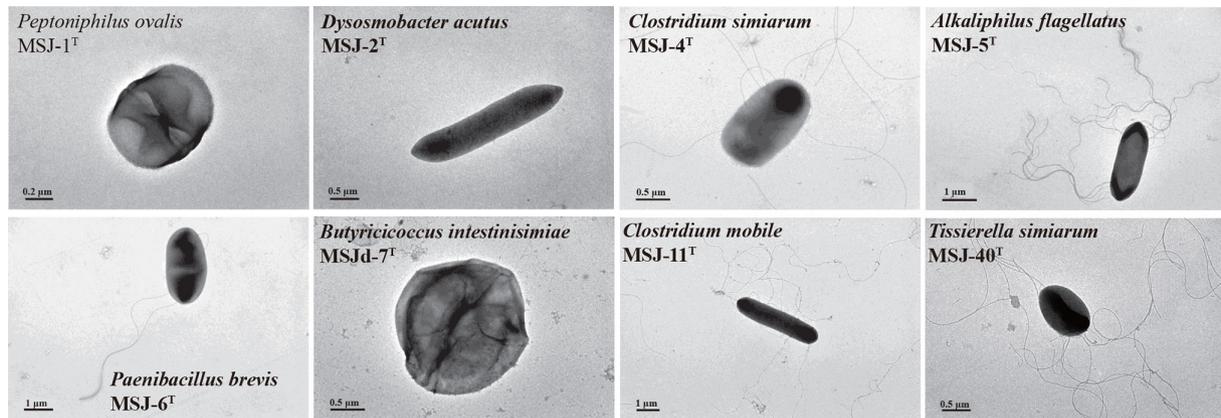


Fig. 1. Cellular morphology (transmission electron microscopy) of eight bacteria isolated from *M. fascicularis*. The names of each bacterium and the size of bars are showed in each picture.

Bacterial growth and cell morphology

Strains MSJ-1^T, MSJ-2^T, MSJ-4^T, MSJ-5^T, MSJ-6^T, MSJd-7^T, MSJ-11^T and MSJ-40^T were strictly anaerobic bacteria. They grew on mGAM agar at 37 °C and formed visible colonies after 1–10 days. Colonies of MSJ-2^T, MSJ-4^T, MSJd-7^T and MSJ-40^T were white, colonies of MSJ-5, MSJ-6 and MSJ-11 were grey, and colonies of MSJ-2 were tiny and translucent. No pigments were observed. Additional features are provided in the species description. Cellular morphology was examined with transmission electron microscopy and are shown in Fig. 1. Cells of strains MSJ-1^T and MSJd-7^T were spherical-shaped. Cells of strains MSJ-2^T, MSJ-4^T, MSJ-5^T, MSJ-6^T, MSJ-11^T and MSJ-40^T were rod-shaped. Flagella were observed for strains MSJ-4^T, MSJ-5^T, MSJ-6^T, MSJ-11^T and MSJ-40^T, but not for MSJ-1^T and MSJd-7^T. Additional features of those strains are detailed in the species description.

Assimilation of carbon sources and fermentative production of SCFAs

The assimilation of 95 carbon sources were tested with Biolog AN MicroPlates and the results are recorded in Fig. 2(a). The eight bacteria showed different carbon source spectra, and in total 75 out of the 95 carbon sources were metabolized. We found that mono- and di-saccharides were preferred by the strains, which are extensively found in guts [50]. The eight bacteria all assimilated five carbon sources, i.e. D-fructose, L-fucose, D-galacturonic acid, palatinose and pyruvic acid.

Many gut microbes produce SCFAs that are related to host health [50, 51]. We determined the production of SCFAs in an mGAM broth that contained glucose, D-galactose, cellobiose, mannose, fructose, rhamnose, palatinose and inulin. The results showed that each strain produced unique profiles of SCFAs (Fig. 2b). Butyric acid was produced by MSJ-1^T, MSJ-4^T, MSJ-5^T, MSJ-6^T, MSJ-11^T and MSJ-40^T. Propionic acid was produced MSJ-4^T, MSJ-11^T, MSJ-6^T and MSJ-40^T. Acetic acid was produced by MSJ-1^T, MSJ-2^T, MSJ-4^T, MSJ-5^T, MSJ-6^T, MSJ-11^T and MSJ-40^T. In addition to the above SCFAs, strains MSJ-6^T and MSJ-11^T produced also branched SCFAs of isobutyric acid and/or isovaleric acid. Strain MSJd-7^T did not produce the six SCFAs detected in this study.

Cellular fatty acid and polar lipid profiling

The chemotaxonomic cellular fatty acid and polar lipid profiles for the eight bacteria were determined and are summarized. As shown in Tables 1 and 2, the eight bacteria had different cellular fatty acid profiles, but they all had C_{14:0}, C_{16:0}, C_{18:0} and anteiso-C_{15:0}. Taking 10% as cutoff value for predominant cellular fatty acids, MSJ-1^T had C_{16:0} (19.9%), MSJ-2^T had C_{16:0} (20.1%), MSJ-4^T had C_{14:0} (15.2%) and C_{16:0} (24.5%), MSJ-5^T had C_{16:0} (10.6%) and iso-C_{13:0} (11.3%)/anteiso-C_{17:0} (14.5%)/anteiso-C_{15:0} (15.7%)/iso-C_{16:0} (16.8%), MSJ-6^T had C_{16:0} (20.6%)/C_{18:0} (10.5%) and iso-C_{16:0} (17.1%)/anteiso-C_{15:0} (25.6%), MSJd-7^T had C_{14:0} (10.9%)/C_{16:0} (26.8%)/C_{18:0} (22.0%) and iso-C_{17:1} ω5c (12.8%), MSJ-11^T had C_{14:0} (19.8%)/C_{16:0} (37.4%)/C_{18:0} (11.7%), and MSJ-40^T had mainly iso-C_{15:0} (62.2%). Polar lipid profiling showed that all eight bacteria had diphosphatidylglycerol and phosphatidylglycerol, but were different from each other in terms of the presence or not of phosphatidylethanolamine, phosphatidylmethylethanolamine, unknown phospholipids, unknown lipids and unknown glycolipids, as detailed in Tables 1 and 2, and Fig. S1 (available in the online version of this article).

General features of genome and genomic DNA G+C contents

The eight bacteria had different genome sizes, MSJ-1^T was the smallest (2.1 Mbp) and MSJ-6^T was the largest genome (5.2 Mbp). Genome sequencing data and some basic features of coding density and G+C contents are listed in Table 3. The genomic DNA

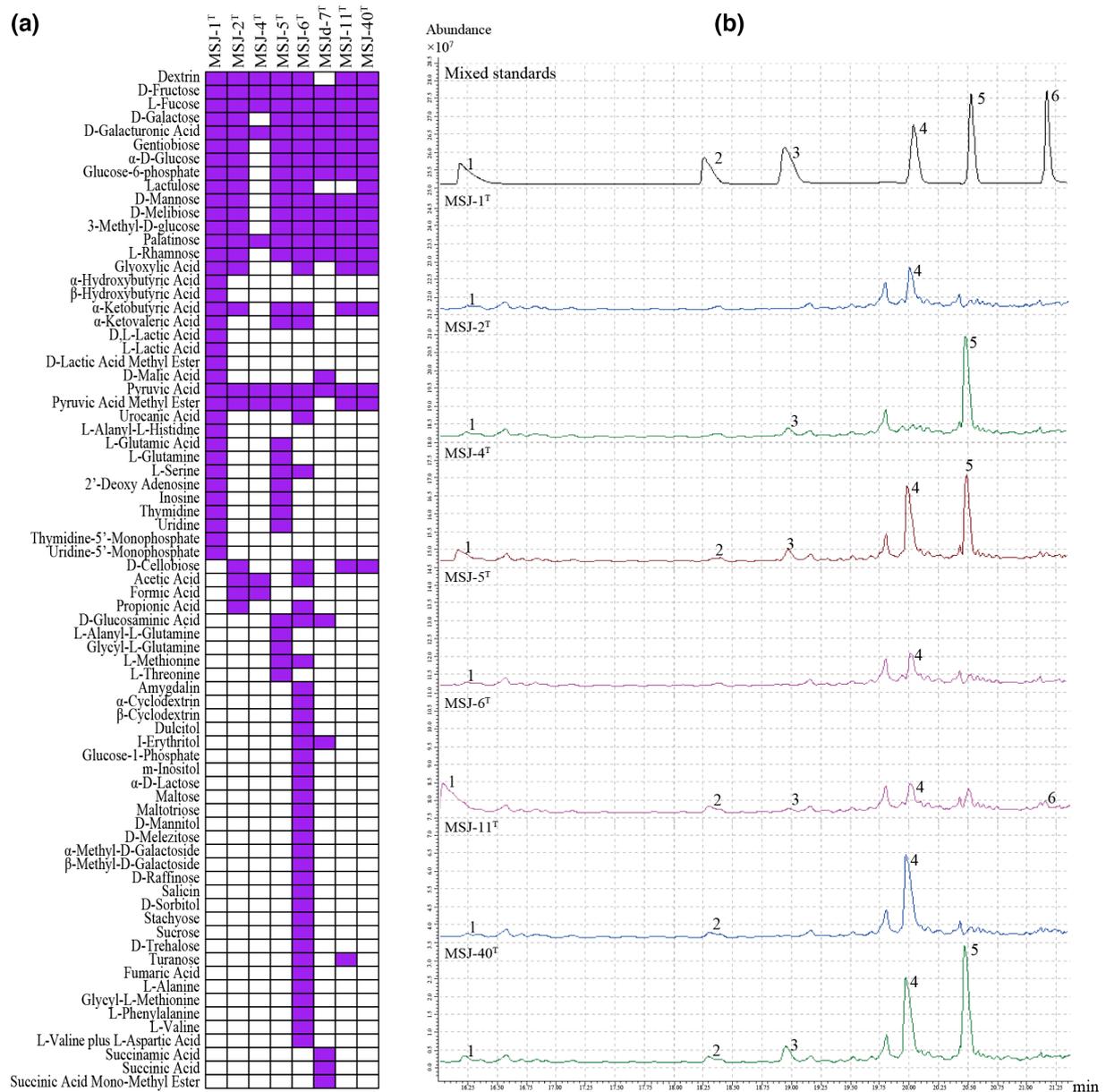


Fig. 2. Assimilation of the carbon sources on Biolog AN MicroPlates (a, purple indicates positive and white indicates negative) and production of short-chain fatty acids in mGAM medium (b). The mixed standards contain six short-chain fatty acids. 1, Acetic acid; 2, propanoic acid; 3, isobutyric acid; 4, butyric acid; 5, isovaleric acid; 6, valeric acid.

G+C molar contents of MSJ-1^T, MSJ-2^T, MSJ-4^T, MSJ-5^T, MSJ-6^T, MSJd-7^T, MSJ-11^T and MSJ-40^T were 30.65, 58.27, 30.46, 31.71, 49.3, 50.29, 30.38, 44.49 and 30.39 mol%, respectively.

The eight bacterial strains represent novel taxa

Based on the 16S rRNA gene and genomic data, we further investigated the phylogenetic and phylogenomic relationships of the eight bacteria to their closely related and currently validly nominated bacterial taxa (Figs 3, S2 and S3). ANI scores based on whole genomes were used to generate UPGMA dendrogram trees (Fig. 4). Combining the results from DNA molecule analysis and the phenotypic characterization, we concluded that MSJ-1^T, MSJ-2^T, MSJ-4^T, MSJ-5^T, MSJ-6^T, MSJd-7^T, MSJ-11^T and MSJ-40^T represented novel species of the currently known genera (for details, see the following sections).

Table 1. Cellular fatty acids and polar acid compositions of strains MSJ-1^T, MSJ-4^T, MSJ-11^T, MSJ-5^T, MSJ-40^T and closely related type strains of the families *Peptoniphilaceae*, *Clostridiaceae* and *Tissierellaceae*

Strains: 1, MSJ-1^T; 2, *P. asaccharolyticus* CCUG 9988^T [55]; 3, *P. gorbachii* WAL 10408^T [55]; 4, MSJ-4^T; 5, MSJ-11^T; 6, *C. liquoris* DSM 100320^T [20]; 7, *C. lundense* DSM 4474^T [20]; 8, MSJ-5^T; 9, *A. halophilus* CGMCC 1.5124^T [63]; 10, *A. oremlandii* DSM 21761^T [63]; 11, MSJ-40^T; 12, *T. praeacuta* DSM 18095^T [27]. Major fatty acid components are indicated with bold text. –, Not detected; TR, trace amount (<1.0%).

Strains	1	2	3	4	5	6	7	8	9	10	11	12
Fatty acids (>10% of total fatty acids):												
C_{16:0}	19.9	14.4	24.0	24.5	37.4	17.7	19.4	10.6	5.0	8.5	7.2	20.0
C _{18:0}	6.8	–	–	3.8	11.7	13.1	3.1	7.7	1.3	4.4	5.1	1.5
C _{14:0}	4.9	5.4	2.9	15.2	19.8	10.8	43.4	5.4	3.2	4.9	9.0	8.0
C _{18:1 ω9c}	4.2	20.2	22.6	1.1	2.22	1.3	–	–	4.0	4.2	–	–
anteiso-C _{15:0}	4.0	–	–	4.1	2.1	–	–	15.7	1.4	3.3	2.3	–
anteiso-C _{17:0}	3.1	1.6	–	0.9	1.0	1.3	1.5	14.5	–	–	0.8	–
iso-C _{16:0}	2.1	–	–	0.9	–	–	–	16.8	–	–	–	TR
iso-C _{15:0}	1.2	2.6	–	9.8	1.1	–	–	9.6	41.4	27.9	62.2	44.5
iso-C _{13:0}	–	–	–	9.3	–	–	–	11.3	9.7	10.5	4.5	–
C _{18:2 ω6,9c} / ante-C _{18:0}	3.6	22.0	21.1	0.6	–	–	–	–	–	–	–	–
Polar lipids*	DPG, PG, PE, PL1, PL2, PL3, L	–	–	DPG, PG, PE, PL1, PL2, PL3, L	DPG, PG, PE, PL1, PL2, PL3, L, APL	PG, PE, PL	–	DPG, PG, PE, PME, PL, L	–	–	DPG, PG, GL1, GL2, GL3, PL1, PL2	–

*APL, aminophospholipid; DPG, diphosphatidylglycerol; PG, phosphatidylglycerol; PE, phosphatidylethanolamine; PME, phosphatidylmethylethanolamine; PL, unknown phospholipids; L, unknown lipid; GL, unknown glycolipid.

Strain MSJ-1^T

The phylogenetic trees revealed that strain MSJ-1^T clustered within the previously described genus *Peptoniphilus* clade, supported by a 100% bootstrap value (Fig. 3d, S2d and S3d). Strain MSJ-1^T was closely related to *Peptoniphilus gorbachii* WAL10418^T (96.82% identity), *Peptoniphilus lacydonensis* DSM 100661^T (95.93%) and *Peptoniphilus harei* DSM 10020^T (95.64%) [24, 52, 53]. In the phylogenomic tree, strain MSJ-1^T also formed a separate branch located in the genus *Peptoniphilus* clade (Fig. 4). The ANI and dDDH values between strain MSJ-1^T and its closest neighbour *P. lacydonensis* DSM 100661^T (GCA 900106515.1) were 78.79 and 20.50%, respectively (Fig. 4). At the time of writing, the genus *Peptoniphilus* contains 20 species with validly published names [54]. Cells of *Peptoniphilus* members are non-spore-forming, obligately anaerobic and coccus-shaped. In addition to its unique 16S rRNA gene and genome sequences, MSJ-1^T contains the predominant fatty acid of C_{16:0}, which is consistent with most species of the genus *Peptoniphilus* (Table 1), although the fatty acids of other members in the genus *Peptoniphilus* are more diverse [55]. Based on phenotypic, chemotaxonomic and phylogenetic results, as well as phylogenomic and genome data, we suggest that strain MSJ-1^T represents a novel species affiliated with the genus *Peptoniphilus* and the name *Peptoniphilus ovalis* sp. nov. is proposed.

DESCRIPTION OF PEPTONIPHILUS OVALIS SP. NOV.

Peptoniphilus ovalis sp. nov. (o.va'lis. L. masc. adj. *ovalis*, pertaining to an egg, egg-shaped).

Cells are non-mobile cocci with diameters of approximately 0.6–0.8 μm, and no flagellum. Strictly anaerobic, heterotrophic growth at 37 °C and pH 7.0. Produces butyric acid and acetic acid from fermentation. After 2 days of cultivation on mGAM agar plate, colonies are 1–2 mm in diameter, white, circular, entire, opaque and smooth. Cells metabolize dextrin, D-fructose, L-fucose, D-galactose, D-galacturonic acid, gentiobiose, α-D-glucose, glucose-6-phosphate, lactulose, D-mannose, melibiose, 3-methyl-D-glucose, palatinose, L-rhamnose, glyoxylic acid, α-hydroxybutyric, β-hydroxybutyric, α-ketobutyric acid, α-ketovaleric acid, D- and L-lactic acid, D-lactic acid methyl ester, D-malic acid, pyruvic acid, pyruvic acid methyl ester, urocanic acid, L-alanyl-L-histidine, L-glutamic acid, L-glutamine, L-serine, 2'-deoxy adenosine, inosine, thymidine, uridine, thymidine-5'-monophosphate and uridine-5'-monophosphate. The predominant cellular fatty acid is C_{16:0}. The polar lipids are diphosphatidylglycerol, phosphatidylglycerol,

Table 2. Cellular fatty acids and polar acids compositions of strains MSJ-2^T, MSJ-6^T, MSJd-7^T and closely related type strains of the families *Oscillospiraceae* and *Paenibacillaceae*

Strain: 1, MSJ-2^T; 2, *D. welbionis* DSM 106889^T [25]; 3, *O. valericigenes* DSM 18026^T [25]; 4, MSJ-6^T; 5, *P. apis* JCM 31620^T [65]; 6, *P. puldeungensis* DSM 27603^T [65]; 7, MSJd-7^T; 8, *B. porcorum* ATCC TSD-102^T [66]; 9, *B. pullicaecorum* DSM 23266^T [66]. Major fatty acid components are indicated with bold text. –, Not detected; TR, trace amount (<1.0%).

Strains	1	2	3	4	5	6	7	8	9
Fatty acids (>10% of total fatty acids):									
C _{16:0}	20.1	TR	14.3	20.6	29.7	7.3	26.8	10.2	–
C _{18:0}	8.7	TR	1.7	10.5	2.4	1.7	22.0	2.3	1.5
C _{14:0}	7.7	2.4	11.5	5.2	3.4	1.5	10.9	15.7	6.4
C _{18:1} ω9c	4.5	TR	–	–	–	14.7	2.51	14.7	–
anteiso-C _{15:0}	8.5	15.2	3.0	25.6	16.7	53.4	4.6	–	–
iso-C _{16:0}	–	TR	–	17.1	3.5	18.9	7.3	–	–
iso-C _{15:0}	3.4	24.2	8.3	8.3	3.9	3.3	–	2.3	1.5
iso-C _{13:0}	1.6	TR	11.8	–	–	–	–	1.6	13.8
iso-C _{17:1} ω5c	–	–	–	–	–	–	12.8	–	–
Polar lipids	DPG, PG, GL1, GL2, GL3, PL	–	–	DPG, PG, PE, GL1, GL2, L	DPG, PG, PE, PL, APL, L	DPG, PG, PE, APL	DPG, PG, PL1, PL2, PL3, GL1, GL2, GL3, L1, L2	–	–

*APL, aminophospholipid; DPG, diphosphatidylglycerol; PG, phosphatidylglycerol; PE, phosphatidylethanolamine; PL, unknown phospholipids; L, unknown lipid; GL, unknown glycolipid.

phosphatidylethanolamine, three unknown phospholipids and an unknown lipid. Genome size is 2102036bp and the G+C content is 30.65 mol%.

The type strain is MSJ-1^T (=CGMCC 1.31770^T=KCTC 15977^T) and was isolated from faecal samples of *M. fascicularis*.

Strain MSJ-2^T

MSJ-2^T was closely related to *Dysosmobacter welbionis* DSM 106889^T (95.78% 16S rRNA gene identity) [25], *Oscillibacter valericigenes* DSM 18026^T (95.15%) and *Oscillibacter ruminantium* JCM 18333^T (94.72%) [56, 57]. The phylogenetic tree indicated that MSJ-2^T formed a heterogeneous cluster with members of *Dysosmobacter* and *Oscillibacter* (Figs 3b, S2b and S3b). The OrthoANI tree clearly separated MSJ-2^T and *D. welbionis* from *O. valericigenes* and *O. ruminantium* (Fig. 4). Furthermore, the G+C molar content of strain MSJ-2^T (58.27 mol%) was closer to that of *D. welbionis* (58.9mol%) than to *O. valericigenes* (53.2%) and *O. ruminantium* (55.0%). Thus, MSJ-2^T was more likely a member of genus *Dysosmobacter*. ANI and dDDH values between strain MSJ-2^T and the closest cultivated neighbour *D. welbionis* DSM 106889^T (GCA 005121165.1) were 74.09 and 21.20%, respectively (Fig. 4), suggesting they represented different species within the genus *Dysosmobacter*. Cells of strain MSJ-2^T were long rods,

Table 3. Genome features of the eight bacterial strains from *M. fascicularis*

Genome features	MSJ-1 ^T	MSJ-2 ^T	MSJ-4 ^T	MSJ-5 ^T	MSJ-6 ^T	MSJd-7 ^T	MSJ-11 ^T	MSJ-40 ^T
Genome Size (bp)	2 102 036	3 161 374	3 811 517	3 614 516	5 239 947	2 711 934	4 014 245	4 088 863
G+C content (mol%)	30.65	58.27	30.46	31.71	49.3	50.29	30.38	30.39
Completeness (%)	98.6	94.63	99.19	97.87	98.66	99.33	100	99.13
Contamination (%)	0.7	0.67	0.93	0.24	0	0.67	0.57	1.98
Number of contigs	14	2	70	51	64	27	18	53
N50 of contigs (bp)	300 523	3 156 307	406 816	534 169	198 601	357 991	1 137 742	199 075
Gene number	2037	3296	3532	3591	4800	3023	4017	4177

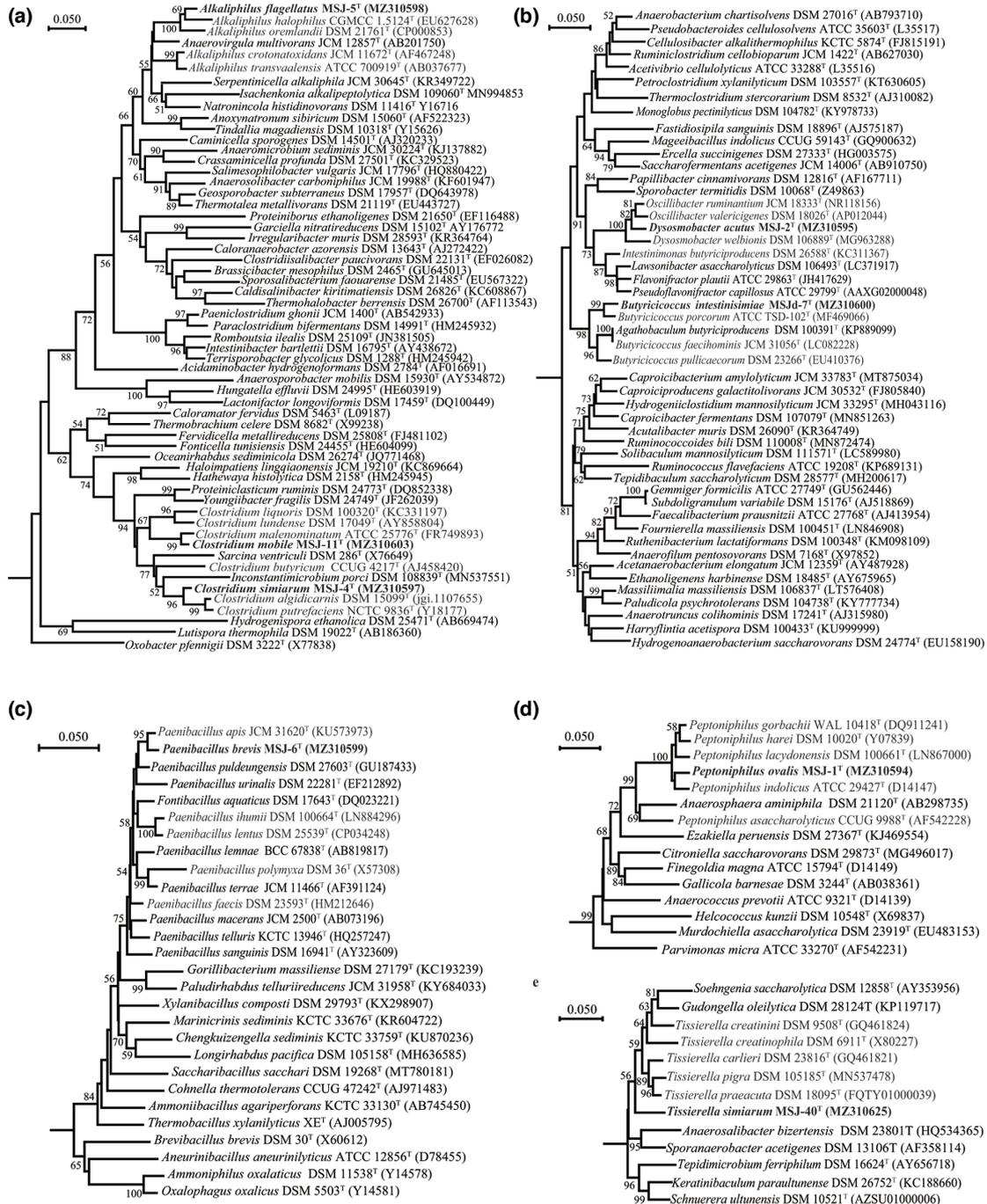


Fig. 3. Phylogenetic trees of the eight bacteria strains based on 16S rRNA gene sequences and using the neighbour-joining algorithm showing the relationships between the eight novel bacteria and their closely related micro-organisms. (a) The phylogenetic tree of strains MSJ-5^T, MSJ-4^T, MSJ-11^T, closely related species in the genera *Alkaliphilus* and *Clostridium*, and the type species of the other genera in the family Clostridiaceae. (b) The phylogenetic tree of strains MSJ-2^T, MSJd-7^T, closely related species in the genera *Dysosmobacter* and *Butyricoccus*, and the type species of the other genera in the family Oscillospiraceae. (c) The phylogenetic tree of strain MSJ-6^T, closely related species in the genus *Paenibacillus*, and the type species of the other genera in the family Paenibacillaceae. (d) The phylogenetic tree of strain MSJ-40^T, closely related species in genus *Tissierella*, and the type species of the other genera in the family Tissierellaceae. (e) The phylogenetic tree of strain MSJ-1^T, closely related species in the genus *Peptoniphilus*, and the type species of the other genera in the family Peptoniphilaceae. GenBank accession numbers are given in parentheses. Bootstrap percentages (>50%) based on 1000 replicates are shown at the nodes. Phylogenetic trees based on the maximum-likelihood method (Fig. S2) and the maximum-parsimony method (Fig. S3) are available as supplementary materials with the online version. *Verrucomicrobium spinosum* DSM 4136 (X90515) was used as an outgroup. Bar, 0.05 substitutions per nucleotide position.

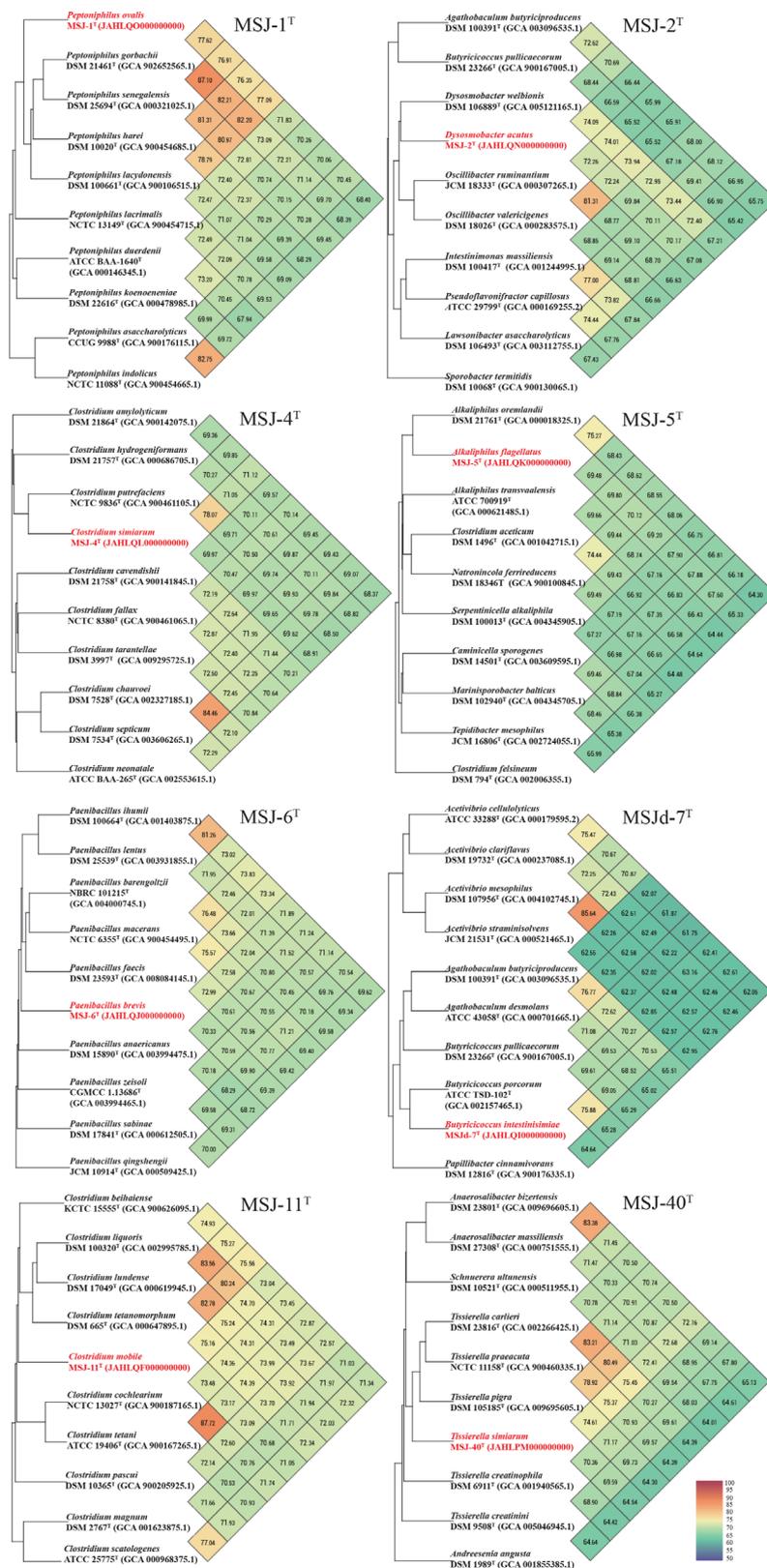


Fig. 4. UPGMA phylogenetic trees and ANI heat maps based on whole genomes. Each of the eight UPGMA phylogenetic trees and the ANI heat maps displays the connections between a novel bacterial taxon and its closely related neighbours, the new names proposed in this study are showed in red. GenBank accession numbers of the genomes are shown in parentheses.

with no flagella. The predominant fatty acid of MSJ-2^T was C_{16:0} (20.1%), which distinguished this organism from *D. welbionis* DSM 106889^T (C_{16:0} <1%) [25] (Table 2). Based on the chemotaxonomic, phylogenetic and genomic results described above, we conclude that strain MSJ-2^T represents a novel species affiliated to the genus *Dysosmobacter* and the name *Dysosmobacter acutus* sp. nov. is proposed.

DESCRIPTION OF *DYSOSMOBACTER ACUTUS* SP. NOV.

Dysosmobacter acutus (a.cu'tus. L. masc. adj. *acutus*, sharp, pointed referring to atypical cell shape).

Cells are non-mobile long rods with sharp ends. No flagellum. The cell size is approximately 0.5–0.6×2.7–2.9 μm. Strictly anaerobic, heterotrophic growth at 37 °C and pH 7.0. Colonies are <1 mm in diameter after 5 days of incubation at 37 °C on mGAM agar plates and flat, circular, entire, translucent and smooth. Fermentative production is isovaleric acid, isobutyric acid and acetic acid. Cells metabolize cellobiose, dextrin, D-fructose, L-fucose, D-galactose, D-galacturonic acid, gentiobiose, α-D-glucose, glucose-6-phosphate, lactulose, D-mannose, melibiose, 3-methyl-D-glucose, palatinose, L-rhamnose, acetic acid, formic acid, glyoxylic acid, α-ketobutyric acid, propionic acid, pyruvic acid and pyruvic acid methyl ester. The predominant cellular fatty acid is C_{16:0}. The major lipids are phosphatidylglycerol and two unknown glycolipids. Genome size is 3161374 bp and the G+C content is 58.27 mol%.

The type strain is MSJ-2^T (=CGMCC 1.32896^T=KCTC 15976^T), which was isolated from a faecal sample of *M. fascicularis*.

Strains MSJ-4^T and MSJ-11^T

Based on the phylogenetic and phynogenomic trees (Figs 3 and 4, S2a, S3a), strains MSJ-4^T and MSJ-11^T formed a cluster that was well-separated from the valid members of the genus *Clostridium* [21]. MSJ-4^T was closely related to *Clostridium algidicarnis* DSM 15099^T (96.85% 16S rRNA gene identity) and *Clostridium putrefaciens* NCTC 9836^T (96.78%) [58, 59] MSJ-11^T was closely related to *Clostridium malenominatum* ATCC 25776^T (98.33%) [60]. The ANI and dDDH values of strain MSJ-4^T to its closest neighbour *Clostridium putrefaciens* NCTC 9836^T (GCA900461105.1) were 78.07 and 22.10%, respectively. The ANI and dDDH values of strain MSJ-11^T to its closest cultivated neighbour *Clostridium beihaiense* KCTC 15555 T (GCA 900626095.1) were 74.93 and 18.50%, respectively. Our results revealed that the major fatty acids of strains MSJ-4^T and MSJ-11^T were C_{16:0} and C_{14:0} which is consistent with the majority of species within genus *Clostridium* [61, 62] (Table 1). Strains MSJ-4^T and MSJ-11^T belonged to genus *Clostridium* and could be differentiated from each other and from other species of genus *Clostridium*. Therefore, we conclude that strain MSJ-4^T represents a novel species and the name *Clostridium simiarum* sp. nov. is proposed, and that strain MSJ-11^T also represents a novel species for which the name *Clostridium mobile* sp. nov. is proposed.

DESCRIPTION OF *CLOSTRIDIUM SIMIARUM* SP. NOV.

Clostridium simiarum (si.mi.a'rum. L. gen. pl. n. *simiarum*, of monkeys).

Cells are fat rods with blunt ends, 0.5–0.9×1.4–2.0 μm, and have peritrichous flagella. Strictly anaerobic, heterotrophic growth at 37 °C and pH 7.0. Produces white, flat, circular, entire, opaque, smooth colonies with a diameter of 2–3 mm after 2 days of incubation at 37 °C on mGAM agar plates. Cells produce butyric acid, propanoic acid and acetic acid, isovaleric acid and isobutyric acid during fermentation. Assimilates the following carbon sources: dextrin, D-fructose, L-fucose, D-galacturonic acid, palatinose, acetic acid, formic acid, pyruvic acid and pyruvic acid methyl ester. The predominant cellular fatty acids are C_{16:0} and C_{14:0}. The polar lipids are diphosphatidylglycerol, phosphatidylglycerol, phosphatidylethanolamine, three unknown phospholipids and an unidentified lipid. Genome size is 3811517 bp and the G+C content is 30.46 mol%.

The type strain is MSJ-4^T (=CGMCC 1.45006^T=KCTC 15975^T) and was isolated from faecal samples of *M. fascicularis*.

DESCRIPTION OF *CLOSTRIDIUM MOBILE* SP. NOV.

Clostridium mobile (mo'bi.le. L. neut. adj. *mobile*, motile).

Cells are rods with size of approximately 0.4–0.7×2.9–9.9 μm, and have flagella at both ends. Strictly anaerobic, with growth at 37 °C and pH 7.0. Colonies are grey, convex, circular, entire, and opaque with a diameter of 2–3 mm after 2 days of incubation at 37 °C on mGAM agar plates. The SCFAs produced by anaerobic fermentation are butyric acid, propanoic acid and acetic acid. Cells metabolize cellobiose, dextrin, D-fructose, L-fucose, D-galactose, D-galacturonic acid, gentiobiose, α-D-glucose, glucose-6-phosphate, D-mannose, melibiose, 3-methyl-D-glucose, palatinose, L-rhamnose, turanose, glyoxylic acid, α-ketobutyric acid, pyruvic acid and pyruvic acid methyl ester. The predominant cellular fatty acids are C_{16:0}, C_{14:0} and C_{18:0}. The polar lipids are diphosphatidylglycerol, phosphatidylglycerol, phosphatidylethanolamine, unidentified aminophospholipid, three unknown phospholipids and an unidentified lipid. Genome size is 4014245 bp and the G+C content is 30.38 mol%.

The type strain is MSJ-11^T (=CGMCC 1.45009^T=KCTC 25065^T) and was isolated from faecal samples of *M. fascicularis*.

Strain MSJ-5^T

Strain MSJ-5^T was closely related to *Alkaliphilus halophilus* CGMCC 1.5124^T (96.29% 16S rRNA gene identity) and *Alkaliphilus oremlandii* DSM 21761^T (96.14%) [63, 64] (Figs 3a, S2a and S3a). Phylogenetic and phylogenomic trees revealed that strain MSJ-5^T was in the genus *Alkaliphilus* clade. The ANI and dDDH values of strain MSJ-5^T to its closest neighbour *A. oremlandii* DSM 21761^T (GCA 000018325.1) were 75.27 and 21.20%, respectively (Fig. 4). Strain MSJ-5^T had rod-shaped and motile cells, which was consistent with the description of the genus *Alkaliphilus*, and had G+C content of 31.71 mol%, within the range of 28–36 mol% for the genus *Alkaliphilus* [63]. The predominant fatty acid compositions vary among *Alkaliphilus* species but iso-C_{15:0}, iso-C_{13:0}, C_{16:0} and C_{14:0} are the major components in most of species, as detected in strain MSJ-5^T. Anteiso-C_{15:0} and anteiso-C_{17:0} were detected in strain MSJ-5^T and distinguish this isolate from other *Alkaliphilus* species (Table 1). Based on the polyphasic analysis, strain MSJ-5^T should be classified as representing a novel species of the genus *Alkaliphilus* for which the name *Alkaliphilus flagellatus* sp. nov. is proposed.

DESCRIPTION OF ALKALIPHILUS FLAGELLATUS SP. NOV.

Alkaliphilus flagellatus (fla.gel.la'tus. L. neut. n. *flagellum*, a whip; L. part. adj. *flagellatus*, flagellated).

Cells are rods, 0.7–1.0×2.2–4.2 μm, and have bundled flagella at both ends. Strictly anaerobic, heterotrophic growth at 37 °C and pH 7.0. Produces grey, low convex, circular, entire, opaque colonies with a diameter of 1–2 mm after 2 days of incubation at 37 °C on mGAM agar plates. Fermentative products are butyric acid and acetic acid. Cells metabolize dextrin, D-fructose, L-fucose, D-galactose, D-galacturonic acid, gentiobiose, D-glucosaminic acid, α-D-glucose, glucose-6-phosphate, lactulose, D-mannose, melibiose, 3-methyl-D-glucose, palatinose, L-rhamnose, α-ketobutyric acid, α-ketovaleric acid, pyruvic acid, pyruvic acid methyl ester, L-alanyl-L-glutamine, L-glutamic acid, L-glutamine, glycyl-L-glutamine, L-methionine, L-serine, L-threonine, 2'-deoxyadenosine, inosine, thymidine and uridine. The predominant cellular fatty acids are iso-C_{16:0}, anteiso-C_{15:0}, anteiso-C_{17:0}, iso-C_{13:0} and C_{16:0}. The polar lipids are diphosphatidylglycerol, phosphatidylglycerol, phosphatidylethanolamine, phosphatidylmethylethanolamine, an unknown phospholipid and an unknown lipid. Genome size is 3614516 bp and the G+C content is 31.71 mol%.

The type strain is MSJ-5^T (=CGMCC 1.45007^T=KCTC 15974^T) and was isolated from faecal samples of *M. fascicularis*.

Strain MSJ-6^T

Strain MSJ-6^T was closely related to *Paenibacillus apis* JCM 31620^T [65], with 96.94% 16S rRNA gene sequence identity. The phylogenetic and phylogenomic analysis showed that strain MSJ-6^T was in the *Paenibacillus* clade (Figs 3 and 4, S2c, S3c). The genome size of MSJ-6^T was 5239947 bp. The ANI and dDDH values of strain MSJ-6^T to its closely related neighbour *Paenibacillus faecis* DSM 23593^T (GCA 008084145.1) were 72.99 and 19.50%, respectively. The unique 16S rRNA and genome sequence is one of the characteristics of MSJ-6^T. The predominant cellular fatty acids of *Paenibacillus* species are anteiso-C_{15:0}, C_{16:0}, iso-C_{16:0} and iso-C_{15:0}. Strain MSJ-6^T shared this profile, but the presence of C_{18:0} differentiates it from other *Paenibacillus* species (Table 2). In addition to diphosphatidylglycerol, phosphatidylglycerol and phosphatidylethanolamine, which is shared by *Paenibacillus* species, MSJ-6^T had two unknown glycolipids (Fig. S1). In terms of phenotypic, chemotaxonomic, phylogenetic and genomic features, strain MSJ-6^T should be classified as representing a novel species of the genus *Paenibacillus* for which the name *Paenibacillus brevis* sp. nov. is proposed.

DESCRIPTION OF PAENIBACILLUS BREVIS SP. NOV.

Paenibacillus brevis sp. nov. (bre'vis. L. masc. adj. *brevis*, short, denoting the formation of short rods).

Cells are ovoid to short rods with size of approximately 0.3–1.42×2.57–3.57 μm, have 1–2 flagella. Strictly anaerobic, heterotrophic growth at 37 °C and pH 7.0. After 5 days of incubation at 37 °C on mGAM agar plates, colonies are 1–2 mm in diameter, grey, circular, entire and translucent. The SCFAs produced by anaerobic fermentation are acetic acid, valeric acid, propanoic acid, butyric acid and isobutyric acid. Cells metabolize amygdalin, cellobiose, α-cyclodextrin, β-cyclodextrin, dextrin, dulcitol, i-erythritol, D-fructose, L-fucose, D-galactose, D-galacturonic acid, gentiobiose, D-glucosaminic acid, α-D-glucose, glucose-1-phosphate, glucose-6-phosphate, m-inositol, lactose, lactulose, maltose, maltotriose, D-mannitol, D-mannose, melezitose, melibiose, 3-methyl-D-glucose, methyl α-D-galactoside, methyl β-D-galactoside, palatinose, raffinose, L-rhamnose, salicin, D-sorbitol, stachyose, sucrose, trehalose, turanose, acetic acid, fumaric acid, glyoxylic acid, α-ketobutyric acid, α-ketovaleric acid, propionic acid, pyruvic acid, pyruvic acid methyl ester, urocanic acid, L-alanine, glycyl-L-methionine, L-methionine, L-phenylalanine, L-serine, L-valine and L-valine plus L-aspartic acid. The predominant cellular fatty acids are anteiso-C_{15:0}, C_{16:0}, iso-C_{16:0} and C_{18:0}. The polar lipids are diphosphatidylglycerol, phosphatidylglycerol, phosphatidylethanolamine, two unknown glycolipids and an unknown lipid. Genome size is 5239947 bp and the G+C content is 49.3 mol%.

The type strain is MSJ-6^T (=CGMCC 1.45008^T=KCTC 15973^T) and was isolated from faecal samples of *M. fascicularis*.

Strain MSJd-7^T

Strain MSJd-7^T was closely related to *Butyricoccus porcorum* ATCC TSD-102^T [66], with 97.15% 16S rRNA gene sequence identity. The genome size of MSJd-7^T is 2711934 bp. The phylogenetic and phylogenomic analysis revealed that strain MSJd-7^T was a member of the genus *Butyricoccus* clade (Figs 3 and 4, S2b, S3b). The ANI and dDDH values of strain MSJd-7^T to its closest neighbour *B. porcorum* ATCC TSD-102^T (GCA 002157465.1) were 75.88 and 21.10%, respectively. *Butyricoccus* species have diverse cellular fatty acid compositions, but the predominant components are C_{14:0}, C_{18:0} and C_{16:0} in most species of the genus. The presence of iso-C_{17:1} ω5c is characteristic of MSJd-7^T and distinguished it from other *Butyricoccus* species (Table 2). According to the phenotypic, chemotaxonomic, phylogenetic and genomic data, strain MSJd-7^T should be classified as representing a novel species of the genus *Butyricoccus* for which the name *Butyricoccus intestinisiae* sp. nov. is proposed.

DESCRIPTION OF *BUTYRICOCOCCUS INTESTINISIMIAE* SP. NOV.

Butyricoccus intestinisiae (in.tes.ti.ni. si'mi.ae. L. neut. n. *intestinum*, intestine; L. fem. n. *simia*, a monkey; N.L. gen. n. *intestinisiae*, of the monkey intestine, where the type strain dwells).

Cells are cocci with a diameter of approximately 1.46–1.85 μm, and no flagellum. Strictly anaerobic, heterotrophic growth at 37 °C and pH 7.0, respectively. Colonies are 2–3 mm in diameter after 2 days of incubation at 37 °C on mGAM agar plates, white, convex, circular, entire, opaque and smooth. Anaerobic and fermentative production of valeric acid. Cells metabolize i-erythritol, D-fructose, L-fucose, D-galactose, D-galacturonic acid, gentiobiose, D-glucosaminic acid, α-D-glucose, glucose-6-phosphate, D-mannose, melibiose, 3-methyl-D-glucose, palatinose, L-rhamnose, D-malic acid, pyruvic acid, succinic acid, succinic acid and succinic acid mono-methyl ester. The predominant cellular fatty acids are C_{16:0}, C_{18:0}, iso-C_{17:1} ω5c and C_{14:0}. The polar lipids are diphosphatidylglycerol, phosphatidylglycerol, four unknown phospholipids, three unknown glycolipids and two unknown lipids. Genome size is 2711934 bp and the G+C content is 50.29 mol%.

The type strain is MSJd-7^T (=CGMCC 1.45013=KCTC 25112) and was isolated from faecal samples of *M. fascicularis*.

Strain MSJ-40^T

Strain MSJ-40^T was closely related to *Tissierella carlieri* DSM 23816^T (94.2% 16S rRNA gene sequence identity), *Tissierella praeacuta* DSM 18095^T (94.13%) and *Tissierella pigra* DSM 105185^T (92.9%) [27, 67, 68]. According to the phylogenetic and phylogenomic trees (Figs 3 and 4, S2e, S3e), MSJ-40^T clustered with members of genus *Tissierella*. Thus, strain MSJ-40^T was likely a member of the genus *Tissierella*. As previously reported for *Tissierella* species [27], cells of strain MSJ-40^T were rod-shaped. The ANI and dDDH values of strain MSJ-40^T to its closest related neighbour *T. pigra* DSM 105185^T (GCA 009695605.1) were 74.61 and 22.40% (Fig. 4). The predominant fatty acid of MSJ-40^T was iso-C_{15:0}, which is a characteristic of the genus *Tissierella* [27, 68] (Table 1). At the time of writing, the genus *Tissierella* has five described species with validly published names, and strain MSJ-40^T is different from them according to phenotypic, chemotaxonomic, phylogenetic and genomic features. Thus, strain MSJ-40^T should be classified as representing a novel species of the genus *Tissierella* and the name *Tissierella simiarum* sp. nov. is proposed.

DESCRIPTION OF *TISSIERELLA SIMIARUM* SP. NOV.

Tissierella simiarum (si.mi.a.rum. L. gen. pl. n. *simiarum*, of monkeys).

Cells are rod-shaped, approximately 0.6–0.8×1.0–3.3 μm, and have flagella at both ends. Strictly anaerobic, heterotrophic growth at 37 °C and pH 7.0. Colonies are 3–5 mm in diameter after 2 days of incubation at 37 °C on mGAM agar plates, white, flat, circular, entire, opaque and smooth. The SCFAs produced by fermentation are isovaleric acid, butyric acid, isobutyric acid, propanoic acid and acetic acid. Cells metabolize cellobiose, dextrin, D-fructose, L-fucose, D-galactose, D-galacturonic acid, gentiobiose, α-D-glucose, glucose-6-phosphate, lactulose, D-mannose, melibiose, 3-methyl-D-glucose, palatinose, L-rhamnose, glyoxylic acid, α-ketobutyric acid, pyruvic acid and pyruvic acid methyl ester. The predominant cellular fatty acid is iso-C_{15:0}. The polar lipids are diphosphatidylglycerol, phosphatidylglycerol, three unknown glycolipids and two unknown phospholipids. Genome size is 4.09 Mb and the G+C content is 30.4 mol%.

The type strain is MSJ-40^T (=CGMCC 1.45012^T=KCTC 25071^T) and was isolated from faecal samples of *M. fascicularis*.

The type strain is MSJ-40^T (=CGMCC 1.45012^T=KCTC 25071^T) and was isolated from faecal samples of *M. fascicularis*.

Funding information

This work was financially supported by the National Key Research and Development Program of China (No.2019YFA0905601), the Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDB38020300), China Microbiome Initiative (CMI) supported by Chinese Academy of Sciences (CAS-CMI).

Acknowledgements

We thank Prof. Yuguang Zhou at Institute of Microbiology, Chinese Academy of Sciences (CAS) for coordination of deposits of type strains. We thank Wenzhao Wang at Institute of Microbiology, Chinese Academy of Sciences (CAS) for the detection of short-chain fatty acids.

Author contributions

Conceived and designed the experiments: S.-J.L., C.L., Q.S. Performed the experiments: D.-H.L., R.A., M.-X.D., Y.-J.W. Sampling: Y.L., P.-J.Y., S.-P.Y. Analysed the data: D.-H.L., C.L., H.-Z.Z., H.-H.C., Z.N., C.-Y.J. Drafted the manuscript: D.-H.L., C.L. Approved final version of manuscript: S.-J.L. All authors read and approved the final manuscript.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Ethical statement

The ethics application (ION-2019043) was approved by the Institute of Neuroscience, Chinese Academy of Sciences.

References

1. Van Treuren W, Dodd D. Microbial contribution to the human metabolome: implications for health and disease. *Annu Rev Pathol* 2020;15:345–369.
2. Liu C, Du M-X, Abuduaini R, Yu H-Y, Li D-H, *et al.* Enlightening the taxonomy darkness of human gut microbiomes with a cultured biobank. *Microbiome* 2021;9:119.
3. Forster SC, Kumar N, Anonye BO, Almeida A, Viciani E, *et al.* A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat Biotechnol* 2019;37:186–192.
4. Zou Y, Xue W, Luo G, Deng Z, Qin P, *et al.* 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nat Biotechnol* 2019;37:179–185.
5. Poyet M, Groussin M, Gibbons SM, Avila-Pacheco J, Jiang X, *et al.* A library of human gut bacterial isolates paired with longitudinal multiomics data enables mechanistic microbiome research. *Nat Med* 2019;25:1442–1452.
6. Browne HP, Forster SC, Anonye BO, Kumar N, Neville BA, *et al.* Culturing of “unculturable” human microbiota reveals novel taxa and extensive sporulation. *Nature* 2016;533:543–546.
7. Lagkouvardos I, Pukall R, Abt B, Foesel BU, Meier-Kolthoff JP, *et al.* Corrigendum: The Mouse Intestinal Bacterial Collection (miBC) provides host-specific insight into cultured diversity and functional potential of the gut microbiota. *Nat Microbiol* 2016;1:16219.
8. Liu C, Zhou N, Du M-X, Sun Y-T, Wang K, *et al.* The Mouse Gut Microbial Biobank expands the coverage of cultured bacteria. *Nat Commun* 2020;11:79.
9. Ghosh S, Pramanik S. Structural diversity, functional aspects and future therapeutic applications of human gut microbiome. *Arch Microbiol* 2021;203:5281–5308.
10. Nagpal R, Wang S, Solberg Woods LC, Seshie O, Chung ST, *et al.* Comparative microbiome signatures and short-chain fatty acids in mouse, rat, non-human primate, and human feces. *Front Microbiol* 2018;9:2897.
11. Manara S, Asnicar F, Beghini F, Bazzani D, Cumbo F, *et al.* Microbial genomes from non-human primate gut metagenomes expand the primate-associated bacterial tree of life with over 1000 novel species. *Genome Biol* 2019;20:299.
12. Podgorski II, Pantó L, Földes K, de Winter I, Jánoska M, *et al.* Adenoviruses of the most ancient primate lineages support the theory on virus-host co-evolution. *Acta Vet Hung* 2018;66:474–487.
13. Li X, Liang S, Xia Z, Qu J, Liu H, *et al.* Establishment of a *Macaca fascicularis* gut microbiome gene catalog and comparison with the human, pig, and mouse gut microbiomes. *Gigascience* 2018;7.
14. Hicks AL, Lee KJ, Couto-Rodríguez M, Patel J, Sinha R, *et al.* Gut microbiomes of wild great apes fluctuate seasonally in response to diet. *Nat Commun* 2018;9:1786.
15. Orkin JD, Campos FA, Myers MS, Cheves Hernandez SE, Guadamuz A, *et al.* Seasonality of the gut microbiota of free-ranging white-faced capuchins in a tropical dry forest. *ISME J* 2019;13:183–196.
16. Newman TM, Shively CA, Register TC, Appt SE, Yadav H, *et al.* Diet, obesity, and the gut microbiome as determinants modulating metabolic outcomes in a non-human primate model. *Microbiome* 2021;9:100.
17. Brinkley AW, Mott GE. Anaerobic fecal bacteria of the baboon. *Appl Environ Microbiol* 1978;36:530–532.
18. Glöckner FO, Yilmaz P, Quast C, Gerken J, Beccati A, *et al.* 25 years of serving the community with ribosomal RNA gene reference databases and tools. *J Biotechnol* 2017;261:169–176.
19. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, *et al.* A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat Biotechnol* 2020;38:1079–1086.
20. Yin Q, Tao Y, Zhu X, Zhou Y, He X, *et al.* *Clostridium liquoris* sp. nov., isolated from a fermentation pit used for the production of Chinese strong-flavoured liquor. *Int J Syst Evol Microbiol* 2016;66:749–754.
21. Collins MD, Lawson PA, Willems A, Cordoba JJ, Fernandez-Garayzabal J, *et al.* The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int J Syst Bacteriol* 1994;44:812–826.
22. Takai K, Moser DP, Onstott TC, Spoelstra N, Pfiffner SM, *et al.* *Alkaliphilus transvaalensis* gen. nov., sp. nov., an extremely alkaliphilic bacterium isolated from a deep South African gold mine. *Int J Syst Evol Microbiol* 2001;51:1245–1256.
23. Ash C, Priest FG, Collins MD. Molecular identification of rRNA group 3 bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test. Proposal for the creation of a new genus *Paenibacillus*. *Antonie van Leeuwenhoek* 1993;64:253–260.
24. Ezaki T, Kawamura Y, Li N, Li ZY, Zhao L, *et al.* Proposal of the genera *Anaerococcus* gen. nov., *Peptoniphilus* gen. nov. and *Gallicola* gen. nov. for members of the genus *Peptostreptococcus*. *Int J Syst Evol Microbiol* 2001;51:1521–1528.
25. Le Roy T, Van der Smissen P, Paquot A, Delzenne N, Muccioli GG, *et al.* *Dysosmobacter welbionis* gen. nov., sp. nov., isolated from human faeces and emended description of the genus *Oscillibacter*. *Int J Syst Evol Microbiol* 2020;70:4851–4858.
26. Eeckhaut V, Van Immerseel F, Teirlinck E, Pasmans F, Fievez V, *et al.* *Butyricoccus pullicaecorum* gen. nov., sp. nov., an anaerobic, butyrate-producing bacterium isolated from the caecal content of a broiler chicken. *Int J Syst Evol Microbiol* 2008;58:2799–2802.
27. Collins MD, Shah HN. NOTES: Reclassification of *Bacteroides praeacutus* Tissier (Holdeman and Moore) in a New Genus, *Tissierella*, as *Tissierella praeacuta* comb. nov. *Int J Syst Bacteriol* 1986;36:461–463.
28. Liu S, Shi W, Zhao G. China microbiome initiative: opportunity and challenges. *Bull Chin Acad Sci* 2017;32:241–250.
29. Lagier J-C, Hugon P, Khelaifia S, Fournier P-E, La Scola B, *et al.* The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clin Microbiol Rev* 2015;28:237–264.
30. Mailhe M, Ricaboni D, Vitton V, Gonzalez J-M, Bachar D, *et al.* Repertoire of the gut microbiota from stomach to colon using culturomics and next-generation sequencing. *BMC Microbiol* 2018;18:157.
31. Duncan SH, Hold GL, Harmsen HJM, Stewart CS, Flint HJ. Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. *Int J Syst Evol Microbiol* 2002;52:2141–2146.
32. Rettedal EA, Gumpert H, Sommer MOA. Cultivation-based multiplex phenotyping of human gut microbiota allows targeted recovery of previously uncultured bacteria. *Nat Commun* 2014;5:4714.
33. Dione N, Khelaifia S, La Scola B, Lagier JC, Raoult D. A quasi-universal medium to break the aerobic/anaerobic bacterial culture dichotomy in clinical microbiology. *Clin Microbiol Infect* 2016;22:53–58.

34. Preston-Mafham J, Boddy L, Randerson PF. Analysis of microbial community functional diversity using sole-carbon-source utilisation profiles - a critique. *FEMS Microbiol Ecol* 2002;42:1–14.
35. Sasser M. Technical note 101: Identification of bacteria by gas chromatography of cellular fatty acids. MIDI. 1990.
36. Minnikin DE, O'Donnell AG, Goodfellow M, Alderson G, Athalye M, et al. An integrated procedure for the extraction of bacterial isoprenoid quinones and polar lipids. *J Microbiol Methods* 1984;2:233–241.
37. Bakir MA, Kitahara M, Sakamoto M, Matsumoto M, Benno Y. *Bacteroides finegoldii* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol* 2006;56:931–935.
38. Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, et al. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 2017;67:1613–1617.
39. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;22:4673–4680.
40. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987;4:406–425.
41. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 1980;16:111–120.
42. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol Biol Evol* 2018;35:1547–1549.
43. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 1993;10:512–526.
44. Fitch WM. Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Biology* 1971;20:406–416.
45. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 1985;39:783–791.
46. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–477.
47. Lee I, Ouk Kim Y, Park SC, Chun J. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 2016;66:1100–1103.
48. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.
49. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;25:1043–1055.
50. Flint HJ, Scott KP, Duncan SH, Louis P, Forano E. Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* 2012;3:289–306.
51. Duncan SH, Flint HJ. Probiotics and prebiotics and health in ageing populations. *Maturitas* 2013;75:44–50.
52. Beyre M, Bakour S, Le Dault E, Rathored J, Michelle C, et al. *Peptoniphilus lacydonensis* sp. nov., a new human-associated species isolated from a patient with chronic refractory sinusitis. *New Microbes New Infect* 2018;23:61–69.
53. Song Y, Liu C, Finegold SM. *Peptoniphilus gorbachii* sp. nov., *Peptoniphilus olsenii* sp. nov., and *Anaerococcus murdochii* sp. nov. isolated from clinical specimens of human origin. *J Clin Microbiol* 2007;45:1746–1752.
54. Parte AC. LPSN-List of Prokaryotic Names with Standing in Nomenclature. *Nucleic Acids Res* 2014;42:D613–D616.
55. Patel NB, Tito RY, Obregón-Tito AJ, O'Neal L, Trujillo-Villaroel O, et al. *Peptoniphilus catoniae* sp. nov., isolated from a human faecal sample from a traditional Peruvian coastal community. *Int J Syst Evol Microbiol* 2016;66:2019–2024.
56. Lee G-H, Rhee M-S, Chang D-H, Lee J, Kim S, et al. *Oscillibacter ruminantium* sp. nov., isolated from the rumen of Korean native cattle. *Int J Syst Evol Microbiol* 2013;63:1942–1946.
57. Iino T, Mori K, Tanaka K, Suzuki KI, Harayama S. *Oscillibacter valerigenes* gen. nov., sp. nov., a valerate-producing anaerobic bacterium isolated from the alimentary canal of a Japanese corbicula clam. *Int J Syst Evol Microbiol* 2007;57:1840–1845.
58. Lawson P, Dainty RH, Kristiansen N, Berg J, Collins MD. Characterization of a psychrotrophic *Clostridium* causing spoilage in vacuum-packed cooked pork: description of *Clostridium algidicarnis* sp. nov. *Lett Appl Microbiol* 1994;19:153–157.
59. Sturges WS, Drake ET. A complete description of *Clostridium putrefaciens* (McBryde). *J Bacteriol* 1927;14:175–179.
60. Breed R, Hitchens AP. *Bergey's Manual of Determinative Bacteriology. Genus II Clostridium*. Baltimore: The Williams & Wilkins Co; 1948, pp. 763–827.
61. Dong Y, Liu Y, Chen N, Zhong Y, Liu L, et al. *Clostridium beihaiense* sp. nov., an anaerobic bacterium isolated from activated sludge. *Int J Syst Evol Microbiol* 2018;68:2789–2793.
62. Zhu H, Fu B, Lu S, Liu H, Liu H. *Clostridium bovifaecis* sp. nov., a novel acetogenic bacterium isolated from cow manure. *Int J Syst Evol Microbiol* 2018;68:2956–2959.
63. Wu X-Y, Shi K-L, Xu X-W, Wu M, Oren A, et al. *Alkaliphilus halophilus* sp. nov., a strictly anaerobic and halophilic bacterium isolated from a saline lake, and emended description of the genus *Alkaliphilus*. *Int J Syst Evol Microbiol* 2010;60:2898–2902.
64. Fisher E, Dawson AM, Polshyna G, Lisak J, Crable B, et al. Transformation of inorganic and organic arsenic by *Alkaliphilus oremlandii* sp. nov. strain OhLAs. *Ann NY Acad Sci* 2008;1125:230–241.
65. Yun JH, Lee JY, Kim PS, Jung MJ, Bae JW. *Paenibacillus apis* sp. nov. and *Paenibacillus intestini* sp. nov., isolated from the intestine of the honey bee *Apis mellifera*. *Int J Syst Evol Microbiol* 2017;67:1918–1924.
66. Trachsel J, Humphrey S, Allen HK. *Butyricoccus porcorum* sp. nov., a butyrate-producing bacterium from swine intestinal tract. *Int J Syst Evol Microbiol* 2018;68:1737–1742.
67. Alauzet C, Marchandin H, Courtin P, Mory F, Lemée L, et al. Multi-locus analysis reveals diversity in the genus *Tissierella*: description of *Tissierella carlieri* sp. nov. in the new class *Tissierellia* classis nov. *Syst Appl Microbiol* 2014;37:23–34.
68. Wylensek D, Hitch TCA, Riedel T, Afrizal A, Kumar N, et al. A collection of bacterial isolates from the pig intestine reveals functional and taxonomic diversity. *Nat Commun* 2020;11:6389.