CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Phylogenetic classification of six novel species belonging to the genus Bifidobacterium comprising Bifidobacterium anseris sp. nov., Bifidobacterium criceti sp. nov., Bifidobacterium imperatoris sp. nov., Bifidobacterium italicum sp. nov., Bifidobacterium margollesii sp. nov. and Bifidobacterium parmae sp. nov
Authors	Lugli, Gabriele A.;Mangifesta, Marta;Duranti, Sabrina;Anzalone, Rosaria;Milani, Christian;Mancabelli, Leonardo;Alessandri, Giulia;Turroni, Francesca;Ossiprandi, Maria Cristina;van Sinderen, Douwe;Ventura, Marco
Publication date	2018-01-01
Original Citation	Lugli, G. A., Mangifesta, M., Duranti, S., Anzalone, R., Milani, C., Mancabelli, L., Alessandri, G., Turroni, F., Ossiprandi, M. C., v Sinderen, D. and Ventura, M. 'Phylogenetic classification of six novel species belonging to the genus Bifidobacterium comprising Bifidobacterium anseris sp. nov., Bifidobacterium criceti sp. nov., Bifidobacterium imperatoris sp. nov., Bifidobacterium italicum sp. nov., Bifidobacterium margollesii sp. nov. and Bifidobacterium parmae sp. nov', Systematic and Applied Microbiology, 41(3), pp.173-183, doi: 10.1016/j.syapm.2018.01.002
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://www.sciencedirect.com/science/article/pii/S0723202018300171 - 10.1016/j.syapm.2018.01.002
Rights	© 2017 Elsevier Ltd. All rights reserved. This manuscript version is made available under the CC BY-NC-ND 4.0 license. - http://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-11-07 02:47:12
Item downloaded from	https://hdl.handle.net/10468/5436



University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Accepted Manuscript



Title: Phylogenetic classification of six novel species belonging to the genus *Bifidobacterium* comprising *Bifidobacterium anseris* sp. nov., *Bifidobacterium criceti* sp. nov., *Bifidobacterium imperatoris* sp. nov., *Bifidobacterium italicum* sp. nov., *Bifidobacterium margollesii* sp. nov. and *Bifidobacterium parmae* sp. nov.

Authors: Gabriele Andrea Lugli, Marta Mangifesta, Sabrina Duranti, Rosaria Anzalone, Christian Milani, Leonardo Mancabelli, Giulia Alessandri, Francesca Turroni, Maria Cristina Ossiprandi, Douwe van Sinderen, Marco Ventura

PII: S0723-2020(18)30017-1
DOI: https://doi.org/10.1016/j.syapm.2018.01.002
Reference: SYAPM 25889

To appear in:

Received date: 9-8-2017
Revised date: 22-12-2017
Accepted date: 8-1-2018

Please cite this article as: Gabriele Andrea Lugli, Marta Mangifesta, Sabrina Duranti, Rosaria Anzalone, Christian Milani, Leonardo Mancabelli, Giulia Alessandri, Francesca Turroni, Maria Cristina Ossiprandi, Douwe van Sinderen, Marco Ventura, Phylogenetic classification of six novel species belonging to the genus Bifidobacterium comprising Bifidobacterium anseris sp.nov., Bifidobacterium criceti sp.nov., Bifidobacterium imperatoris sp.nov., Bifidobacterium italicum sp.nov., Bifidobacterium margollesii sp.nov.and Bifidobacterium parmae sp.nov., Systematic and Applied Microbiology https://doi.org/10.1016/j.syapm.2018.01.002

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Phylogenetic classification of six novel species belonging to the genus *Bifidobacterium* comprising *Bifidobacterium anseris* sp. nov., *Bifidobacterium criceti* sp. nov., *Bifidobacterium imperatoris* sp. nov., *Bifidobacterium italicum* sp. nov., *Bifidobacterium margollesii* sp. nov. and *Bifidobacterium parmae* sp. Nov

Running title: Identification of novel bifidobacterial species

Gabriele Andrea Lugli*[1], Marta Mangifesta*[2], Sabrina Duranti*[1], Rosaria Anzalone[1], Christian Milani[1], Leonardo Mancabelli[1], Giulia Alessandri[1], Francesca Turroni[1], Maria Cristina Ossiprandi[3], Douwe van Sinderen[4] and Marco Ventura[1]

Laboratory of Probiogenomics, Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parma, Italy[1]; GenProbio srl, Parma, Italy[2]; Department of Veterinary Medical Science, University of Parma, Parma, Italy[3]; APC Microbiome Institute and School of Microbiology, National University of Ireland, Cork, Ireland[4].

*These authors contributed equally

Corresponding author. Mailing address for Marco Ventura laboratory of Probiogenomics, Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parco Area delle Scienze 11a, 43124 Parma, Italy. Phone: ++39-521-905666. Fax: ++39-521-905604. E-mail: marco.ventura@unipr.it

ACCEPTED MANUSCRIPT

Abstract

Six *Bifidobacterium* strains, i.e., Goo31D, Ham19E, Rab10A, Tam1G, Uis4E and Uis1B, were isolated from domestic goose (*Anser domesticus*), European hamster (*Cricetus cricetus*), European rabbit (*Oryctolagus cuniculus*), emperor tamarin (*Saguinus imperator*) and pygmy marmoset (*Callithrix pygmaea*), respectively. Cells are Gram-positive, non-motile, non-sporulating, facultative anaerobic and fructose 6-phosphate phosphoketolase-positive. Phylogenetic analyses based on 16S rRNA and ITS sequences, multilocus sequences (including *hsp*60, *rpo*B, *dna*J, *dna*G and *clp*C genes) and the core genome revealed that bifidobacterial strains Goo31D, Ham19E, Rab10A, Tam1G, Uis4E and Uis1B exhibit close phylogenetic relatedness to *Bifidobacterium choerinum* LMG 10510, *Bifidobacterium hapali* DSM 100202, *Bifidobacterium saguini* DSM 23967 and *Bifidobacterium stellenboschense* DSM 23968. Further genotyping based on the genome sequence of the isolated strains combined with phenotypic analyses, clearly show that these strains are distinct from each of the type strains of the so far recognized *Bifidobacterium* species. Thus, *Bifidobacterium anseris* sp. nov. (Goo31D = LMG 30189^T = CCUG 70960^T), *Bifidobacterium criceti* sp. nov. (Ham19E = LMG 30188^T = CCUG 70962^T), *Bifidobacterium imperatoris* sp. nov. (Tam1G = LMG 30297^T = CCUG 70961^T), *Bifidobacterium italicum* sp. nov. (Rab10A = LMG 30187^T = CCUG 70963^T), *Bifidobacterium margollesii* sp. nov. (Uis1B = LMG 30296^T = CCUG 70959^T) and *Bifidobacterium parmae* sp. nov. (Uis4E = LMG 30295^T = CCUG 70964^T) are proposed as novel *Bifidobacterium* species.

Keywords: phylogenetic, next generation sequencing, genomics, metagenomics, bifidobacteria, *Bifidobacterium*

ACCEPTED MANUSCRIPT

Introduction

Bifidobacteria are one of the most noteworthy microbial groups belonging to the *Actinobacteria* phylum, and are common inhabitants of the gastro-intestinal (GIT) of mammals, birds and social insects [1-3]. Notably, this group of microorganisms is believed to confer a range of health-promoting properties to its host, such as modulation of immune response, modulation of mucosal physiology of the host and inhibition of pathogen proliferation [4-6]. Furthermore, bifidobacteria are known to be abundantly present in those animals that provide parental care to their offspring and several species belonging to this genus are reported to be among the first gut colonizers of newborns [7-10]. Currently 59 taxa, representing 50 species and nine subspecies have been formally recognized as members of the genus *Bifidobacterium* [11-18].

A recent metagenomics-based study, based on Internally Transcribed Spacer (ITS) rRNA profiling and aimed at exploring the gut microbiota biodiversity across the mammalian branch of the tree of life, identified a number of putative novel taxa belonging to the genus *Bifidobacterium* [19]. Interestingly, these data revealed the presence of 89 putative novel bifidobacterial taxa in addition to the previously described (sub)species [19]. The 291 analysed animal hosts included the domestic goose (*Anser domesticus*), a winged animal that belongs to the *Anatidae* family, the European hamster (*Cricetus cricetus*), a rodent that belongs to the *Cricetidae* family, the European rabbit (*Oryctolagus cuniculus*), a lagomorph belonging to the *Leporidae* family, and the emperor tamarin (*Saguinus imperator*) and pygmy marmoset (*Callithrix pygmaea*), two primates harboring the *Callitrichidae* family that live in the Amazonian lowland.

In the current study, we describe the identification of novel bifidobacterial species based on 16S rRNA and ITS profiling, followed by genomic comparison as based on whole genome sequencing. Genomic investigation as well as phylogenetic and phenotypic analyses allowed the identification of six proposed novel bifidobacterial species isolated from animal feces, being related to *Bifidobacterium choerinum*

ACCEPTED MANUSCRIPT

LMG 10510, *Bifidobacterium hapali* DSM 100202, *Bifidobacterium tissieri* DSM100201, *Bifidobacterium saguini* DSM 23967 and *Bifidobacterium stellenboschense* DSM 23968.

Materials and Methods

Bifidobacterial selection. Fecal samples of several animals were collected by biologist supervisors of Italian zoo's and Italian Natural Parks in collaboration with the Laboratory of Probiogenomics, University of Parma, Italy. One gram of fecal sample of a domestic goose (*Anser domesticus*), European hamster (*Cricetus cricetus*), European rabbit (*Oryctolagus cuniculus*), Emperor tamarin (*Saguinus imperator*) and Pygmy marmoset (*Callithrix pygmaea*) was mixed with nine ml of phosphate-buffered saline (PBS), pH 6.5. Serial dilution and subsequent plating were performed using de Man-Rogosa-Sharpe (MRS) agar, supplemented with 50 μg/ml mupirocin (Delchimica, Italy) and 0.05 % (wt/col) L-cysteine hydrochloride. Agar plates were incubated for 48 h at 37°C in a chamber (Concept 400; Ruskin) with anaerobic atmosphere (2.99 % H_2, 17.01 % CO_2 and 80 % N_2). Morphologically different colonies that developed on MRS plates were randomly picked and re-streaked in order to isolate purified bacterial strains. All isolates were subjected to DNA isolation and characterized as previously described by Turroni *et al.* [20].

ERIC-PCR genotyping. The isolates were subjected to molecular typing through the use of Enterobacterial Repetitive Intergenic Consensus sequences (ERIC) PCR. ERIC-PCR was carried out using ERIC-1 (5′-ATGTAAGCTCCTGGGGATTCAC-3′) and ERIC-2 (5′-AAGTAAGTGACTGGGGTGAGCG-3′) primers following a previously described procedure [21].

Amplification of 16S rRNA gene and associated ITS sequences. Partial 16S rRNA gene sequences were amplified from extracted DNA using primer pair Probio_Uni/Probio_Rev, which targets the V3

ACCEPTED MANUSCRIPT

region of the 16S rRNA gene sequence [22]. Partial ITS sequences of the six novel bifidobacterial taxa were amplified using primer pair Probio-bif_Uni/Probiobif_Rev [23].

Genome sequencing and assemblies. DNA extracted from the bifidobacterial isolates was subjected to whole genome sequencing using MiSeq (Illumina, UK) at GenProbio srl (Parma, Italy) following the supplier's protocol (Illumina, UK). Fastq files of the paired-end reads obtained from targeted genome sequencing of the isolated strains were used as input for the genome assemblies through the MEGAnnotator pipeline [24]. MIRA (version 4.0.2) software was used for *de novo* assembly of each bifidobacterial genome sequence [25], while protein-encoding open reading frames (ORFs) were predicted using Prodigal [26].

Bifidobacterial core genome evaluation. For the 59 genomes of bifidobacterial type strains and the six novel sequenced genomes, a pan-genome calculation was performed using the PGAP pipeline [27]. The ORF content of each bifidobacterial genome was organized in functional gene clusters using the GF (Gene Family) method involving comparison of each protein to all other proteins using BLAST analysis (cut-off E-value of 1×10^{-5} and 50 % identity over at least 80 % of both protein sequences). Sequences were then clustered in protein families named clusters of orthologous genes (COGs), using MCL (graph-theory-based Markov clustering algorithm) [28] and protein families shared between all genomes, named core COGs, were defined by selecting the families that contained at least one single protein member for each genome.

Phylogenetic and phylogenomic comparisons. The collected bifidobacterial 16S rRNA gene sequences were aligned using MAFFT [29], as well as the ITS, housekeeping genes and core gene sequences. Bifidobacterial phylogenetic trees were constructed using the neighbor-joining method in Clustal W, version 2.1 [30] and were built using FigTree (http://tree.bio.ed.ac.uk/software/figtree/). For each genome pair, a value of average nucleotide identity (ANI) was calculated using the program JSpecies,

ACCEPTED MANUSCRIPT

version 1.2.1 [31]. The identity percentage between 16S rRNA gene or ITS sequences was calculated with MatGat, version 2.03 (Matrix Global Alignment Tool), using the BLOSUM 50 alignment matrix [32]. The Genome-to-Genome Distance Calculator (GGDC) version 2.1 was employed to estimate the DNA-DNA hybridization (DDH) between bifidobacterial taxa, using the recommended "Formula 2" (identities / high-scoring segment pairs length) [33].

Phenotypic characterization. The morphology of six novel bifidobacterial taxa was determined using phase-contrast microscopy after incubation of each strain under anaerobic conditions at 37°C for 24h. Growth was assessed, using MRS broth, under aerobic or anaerobic conditions at different temperatures (20, 25, 30, 35, 37, 40 and 45 °C) and at different pH values (pH 3.5, 4.0, 4.5, 5.0, 5.5 and 6.0) (Table 1).

Gram staining was performed with LIVE *Bac*Light™ Bacterial Gram Stain Kit (ThermoFisher Scientific) as previously described by Duranti *et al.* [11]. The six identified, novel bifidobacterial taxa were grown on MRS medium without glucose supplemented with 1 % (wt/vol) of a particular sugar. Cultures were cultivated in the wells of a 96-well microtiter plate, with each well containing a different sugar, and incubated in an anaerobic cabinet at 37°C. The optical densities at 600nm (OD_{600}) was determined using a plate reader (Biotek, VT, USA) at two different time points, 24 and 48 h. Growth assays were carried out in duplicate and non-inoculated MRS medium was used as a negative control. Carbohydrates tested in this study include arabinose, arabinogalactan, cellobiose, fructo-oligosaccharides, fructose, fucose, galactose, galacto-oligosaccharides, glucose, glycogen, inulin, lactose, maltodextrin, maltose, maltotriose, mannitol, mannose, melibiose, N-acetyl-D-galactosamine, N-acetyl-D-glucosamine, pullulan, raffinose, rhamnose, ribose, sorbitol, starch, sucrose, trehalose, turanose and xylose (Table 2). The six novel isolates and their closest related species, i.e., *B. choerinum* LMG10510, *B. tissieri* DSM100201, *B. hapali* DSM100202, *B. saguini* DSM23967 and *B.*

ACCEPTED MANUSCRIPT

stellenboschense DSM 23968^T, were also investigated for enzymatic activities, using API 50 CHL and Rapid ID 32 test kits (BioMérieux) (Table 2). The cell-wall peptidoglycan composition was analyzed following the protocols previously described by Schumann [34]. Cellular fatty acids methyl esters were determined according to the protocol described by Miller [35] with slight modifications [36]. The results analyzed using MIS Standard Software (Microbial ID) and showed that palmitic, myristic and oleic acid were the dominant fatty acids (Table S1).

Bifidobacterial genome sequences. Whole Genome Shotgun projects were deposited at DDBJ/ENA/GenBank under the accession numbers MVOG00000000, MVOH00000000, NMWT00000000, NMWU00000000, NMWV00000000 and NMYC00000000. The versions described in this paper are accessible under the following numbers: MVOG01000000, MVOH01000000, NMWT01000000, NMWU01000000, NMWV01000000 and NMYC01000000.

Results and Discussion

Novel bifidobacterial strain selection and *in vitro* characterization. Fecal samples were screened in order to assess the bifidobacterial population of the gut environment of a variety of animals, including goose, hamster, rabbit and monkeys (Table 3). A combination of culture-based, molecular-based and genomics-based methods was applied in order to obtain a detailed characterization of each strain. A total of 32 bifidobacterial strains were isolated from stool samples from five different animals, including domestic goose (*Anser domesticus*), European hamster (*Cricetus cricetus*), European rabbit (*Oryctolagus cuniculus*), emperor tamarin (*Saguinus imperator*) and pygmy marmoset (*Callithrix pygmaea*). The genotypic characterization by ERIC-PCR revealed the presence of three known species, *B. choerinum*, *B. saguini* and *B. stellenboschense* by means of comparison to ERIC profiles of the type strains of the respective species (Fig. S1). Furthermore, a representative strain belonging to each putative novel species, was characterized in detail, i.e., strains Goo31D, Ham19E, Rab10A, Tam1G, Uis4E and Uis1B,

ACCEPTED MANUSCRIPT

which correspond to the bifidobacterial taxa previously identified with the names new_taxa_72, new_taxa_71, new_taxa_70, new_taxa_45, new_taxa_77 and new_taxa_61, respectively [19]. These strains represent Gram-positive bacteria, that are all non-motile, non-sporulating and F6PKK-positive. Growth of all isolates and known correlated species was examined on 30 different carbon sources (Table 2). As shown in Table 2, fermentation capabilities for all sugars, except for glucose, appear to be variable among the novel strains. Interestingly, metagenomic analyses performed by Milani *et al* [19]. identified the presence of these new taxa in other animals, i.e. Dolphin (*Delphinus delphis*), European hedgehog (*Erinaceus europaeus*), leopard (*Panthera pardus*) and Norwegian forest cat (*Felis catus*). This suggests that these bifidobacterial taxa can be found in other mammalian species.

Phylogenetic analyses based on 16S rRNA and ITS hypervariable regions. A first characterization of the isolated bifidobacterial taxa was performed through 16S rRNA gene and ITS sequencing, followed by comparison with the same sequenced molecular markers retrieved from public databases. Notably, the ability to distinguish closely related bifidobacterial taxa based on an ITS sequence dataset of 48 bifidobacterial type strains has previously been described [23]. The 16S rRNA gene sequences of Goo31D, Ham19E, Rab10A, Tam1G, Uis4E and Uis1B showed identity values ranging from 96.2 % to 98.6 % with respect to *B. choerinum* LMG 10510, *Bifidobacterium saguini* DSM 23967, *Bifidobacterium stellenboschense* DSM 23968 and *Bifidobacterium tissieri* DSM 100201 (Table 3). Furthermore, the hypervariable ITS sequences of the six isolated strains displayed identity values ranging from 65.7 % to 89 % against *Bifidobacterium biavatii* DSM 23969 and *B. saguini* DSM 23967 (Table 3). Notably, identity values of hypervariable ITS sequences identified among the nine subspecies of the genus *Bifidobacterium* range from 90.4 % to 98.9 % between *Bifidobacterium animalis* and *Bifidobacterium longum* subspecies. Thus, data retrieved from the ITS analysis highlight a high degree of sequence diversity between the isolates from the classified members of the genus *Bifidobacterium*. Therefore, the hypervariable region of the ITS sequence corresponding to the region between primers Probio-

ACCEPTED MANUSCRIPT

bif_Uni/Probio-bif_Rev [23], was employed to build a phylogenetic tree to better assess the phylogenetic relatedness between strains (Fig. S2). In a similar fashion, a phylogenetic tree based on 16S rRNA gene sequences was constructed (Fig. S3). These two phylogenetic trees displayed similar results for four of the six isolated strains. In this context, Goo31D and Rab10A share the same phylogenetic position in both trees with *B. choerinum* LMG 10510, while Tam1G and Uis4E are placed on the same phylogenetic branch together with *B. saguini* DSM 23967 and *B. stellenboschense* DSM 23968, respectively (Figs. S2 and S3). Notably, both Ham19E and Uis1B do not cluster with any bifidobacterial strain in the ITS-based phylogenetic tree. Furthermore, based on the 16S rRNA gene-based tree, both of these isolates are related to *B. choerinum* LMG 10510 and the *B. tissieri* DSM 100201-*Bifidobacterium vansinderenii* LMG 30126 pair [11-18], respectively (Fig. S3).

Phylogenetic analysis based on housekeeping genes. In order to further explore the genetic differences between the six isolated strains and the currently recognized (sub)species of the *Bifidobacterium* genus, nucleotide identity between five genes, i.e. those of *hsp*60, *rpo*B, *dna*J, *dna*G and *clp*C, which are considered to represent molecular clocks in (bifido)bacterial taxonomy, was evaluated [37, 38]. These five selected housekeeping genes showed identity values (at deduced amino acid level) ranging from 82.3 % to 100 % for Uis1B *dna*G and Tam1G *clp*C, with respect to *B. tissieri* DSM 100201 and *B. saguini* DSM 23967 gene sequences (Table S2). Thus, in order to unveil the phylogenetic relatedness between isolates Goo31D, Ham19E, Rab10A, Tam1G, Uis4E and Uis1B on the one hand, and currently recognized (sub)species of the genus *Bifidobacterium* on the other, a phylogenetic tree was constructed based on the amino acid sequences of the above-mentioned house-keeping genes (Fig. S4).

The phylogenetic tree based on the five housekeeping genes displayed a higher robustness compared to that achieved by 16S rRNA gene sequences, as displayed by the bootstrap values (Figs. S3 and S4). While Goo31D, Ham19E and Rab10A displayed genetic relatedness with *B. choerinum* LMG 10510, Tam1G and Uis4E shared the same branches of *B. saguini* DSM 23967 and *B. stellenboschense* DSM

ACCEPTED MANUSCRIPT

23968, respectively. Furthermore, among the so far identified bifidobacterial strains, Uis1B displayed the highest level of housekeeping gene diversity when compared with other bifidobacteria, sharing a phylogenetic position closer to that of *B. tissieri* DSM 100201 and *B. vansinderenii* LMG 30126. Thus, an in depth *in silico* analysis based on genome sequences was performed in order to better classify these novel bifidobacterial isolates.

Phylogenomic characterization of the novel bifidobacterial taxa. To get insights into the genetic similarities between Goo31D, Ham19E, Rab10A, Tam1G, Uis4E and Uis1B, and publicly available genome sequences of the genus *Bifidobacterium*, the six isolated genomes were decoded by means of a Next Generation Sequencing (NGS) approach. The sequenced bifidobacterial strains were each shown to contain a single chromosome ranging in length from 2,155,882 to 2,820,211 for Ham19E and Uis4E, respectively, with an average fold coverage ranging from 55.52 to 139.27 (Table 3). The genome sequences of each isolated taxon were used to assess the genetic similarity at genomic level with respect to other currently recognized bifidobacterial (sub)species by Average Nucleotide Identity (ANI) analysis [39]. The assembled genome of Tam1G exhibits the highest ANI value (94.55 %) against *B. saguini* DSM 23967, while the Uis4E genome sequence is strictly correlated with *B. stellenboschense* DSM 23968 (93.45 %) (Table 3). Furthermore, Uis1B displays the highest ANI value (88.04 %) with respect to *Bifidobacterium hapali* DSM 100202, while Goo31D, Ham19E and Rab10A exhibit ANI values that range from 91.8 % to 87.81 % with respect to *B. choerinum* LMG 10510 (Table 3). In this context, it should be noted that two strains displaying an ANI value <95 % are considered to belong to two distinct species [31]. Notably, these observed ANI values are below the assigned threshold value used for species recognition, supporting the notion that the investigated strains possess a unique genomic composition as compared to currently recognized bifidobacterial taxa (for which a genome sequence is available) [39]. Furthermore, strains Goo31D, Ham19E and Rab10A that exhibit a high level of genomic identity to *B. choerinum* LMG 10510, reveal the highest ANI value of 91.76 % between Ham19E and Rab10A, once

ACCEPTED MANUSCRIPT

again generating an ANI value below 95 %. Besides, Uis1B showed similar ANI values with respect to the phylogenetic correlated taxa *B. tissieri* DSM 100201 (86.81 %) and *B. vansinderenii* LMG 30126 (86.77 %), revealing similar genomic relatedness with three strains isolated from monkeys.

Genome sequencing of Goo31D, Ham19E, Rab10A, Tam1G, Uis4E and Uis1B also allowed us to compare the genetic makeup between these strains and other members of the *Bifidobacterium* genus. A comparative study was undertaken to determine orthology between bifidobacterial strains. Notably, such an *in silico* analysis allows the identification of 259 clusters of orthologous genes (COGs) shared by all bifidobacterial taxa, representing the core genome of currently sequenced *Bifidobacterium* representatives. Of note, conserved genes, which represent paralogs within bifidobacterial genomes, were not considered, and a phylogenetic core genome tree was built based on the concatenation of 233 protein sequences (Fig. 1). The phylogenetic tree based on the core genome confirmed the positioning of the six isolated strains within the *Bifidobacterium* genus that were observed in the phylogenetic analyses based on five housekeeping genes (Fig. S4). Moreover, the subdivision in seven bifidobacterial phylogenetic groups allowed the distribution of the identified strains into two major groups, i.e., *Bifidobacterium pseudolongum* group for Goo31D, Ham19E and Rab10A, and *B. longum* group for Tam1G and Uis4E. In contrast, Uis1B does not belong to any of the so far identified bifidobacterial phylogenetic groups, yet shares a phylogenetic position close to other bifidobacterial strains that have been isolated from monkeys [11], [13].

According to phylogenetic analyses based on ITS, 16S rRNA and housekeeping gene sequences, and phylogenomic analyses based on the bifidobacterial core gene sequences, strains Goo31D, Ham19E, Rab10A, Uis4E and Uis1B are shown to be genetically different from any other bifidobacterial type strain characterized to date. Thus, the novel taxon, *Bifidobacterium anseris* sp. nov., *Bifidobacterium criceti* sp. nov., *Bifidobacterium italicum* sp. nov., *Bifidobacterium parmae* sp. nov. and *Bifidobacterium margollesii* sp. nov., are proposed here. Furthermore, a Genome-to-Genome Distance analysis was

ACCEPTED MANUSCRIPT

performed to examine in depth the genomic relatedness between Tam1G and the phylogenomic correlated strain *B. saguini* DSM 23967 [33]. This analysis estimates a DNA-DNA hybridization (DDH) based on the recommended "Formula 2" (identities / high-scoring segment pairs length) of 55.2 % (confidence interval 52.5 - 57.9 %). The assignment of Tam1G to a novel bifidobacterial species was based on an ANI value below 95 % and an estimated DDH of well below 70 % when compared to its closest relative *B. saguini* DSM 23967, thus corroborating that *Bifidobacterium imperatoris* sp. nov. should be considered a novel bacterial species.

Description of *Bifidobacterium italicum* sp. nov.

Bifidobacterium italicum (i.ta.li.cum, N.L. gen. n. *italicum* pertaining to Italy; country where the type strain was isolated).

Cells are Gram-positive, non-motile, non-sporulating and F6PPK-positive and, when grown in MRS broth under anaerobic condition, showed a small rod morphology (Fig. 2). Cells grow in both anaerobic and aerobic conditions. Colony diameter, when grown on MRS agar, ranges from 1.0 to 2.5 mm. Furthermore, colonies grown on the surface of an MRS agar plate are white and circular, while embedded colonies are white and elliptical. Moreover, after 48h the cells incubated under aerobic and anaerobic conditions are able to grow at temperatures that range from 25 to 42°C, but not at 20ºC or 45ºC (Table 1). Moreover, the cells grow at pH 5-6 (Table 1).

Fermentation profiles of *B. italicum* LMG 30187 revealed that it is able to ferment a wide range of simple and complex carbohydrates, including cellobiose, glucose, maltodextrin, maltose, melibiose and pullulan (i.e., reaching a final OD_{600} of ≥ 0.51). In contrast, *B. italicum* LMG 30187 displays little if any growth on arabinogalactan, fructo-oligosaccharides, galactose, galacto-oligosaccharides, glycogen, inulin, lactose, maltotriose, mannitol, mannose, N-acetyl-D-glucosamine, N-acetyl-D-galactosamine, raffinose, ribose, sorbitol, starch, sucrose, trehalose and turanose (i.e., final OD_{600} is ≤ 0.5) (Table 2). Furthermore,

ACCEPTED MANUSCRIPT

the strain is unable to grow on arabinose, cellobiose, fructose, fucose, rhamnose and xylose. Positive enzymatic activities are observed for α-arabinosidase, β-glucuronidase, glutamic acid decarboxylase, alkaline phosphatase, arginine arylamidase, leucyl-glycine arylamidase, pyroglutamic acid arylamidase, glutamyl glutamic acid and serine arylamidase. In contrast, only weak reactions were observed for α-galactosidase, β-galactosidase and α-glucosidase activities (Table 2). Moreover, N-acetyl-β-glucosaminidase, arginine dihydrolase and β-glucosidase activities were not detected (Table 2).

The peptidoglycan type is L-Lys-L-Ser-L-Ala_2. The major fatty acids identified are palmitic, myristic, oleic acid and stearic acid-dimetylacetal (Table S1).

Phylogenetic analysis of *B. italicum* LMG 30187 highlights that this strain is closely related to *B. choerinum* LMG 10510.

The type strain Rab10A (=LMG 30187^T=CCUG 70963^T) was isolated from a faecal sample of a European rabbit (*Oryctolagus cuniculus*). The DNA G+C content is 65.45 %. Digital Protologue Taxonumber: TA00239.

Description of *Bifidobacterium criceti* sp. nov.

Bifidobacterium criceti (cri.ce.ti, N.L. gen. n. *criceti* of *cricetus*, common scientific name of European hamster, *Cricetus cricetus*).

Cells are Gram-positive, non-motile, non-sporulating and F6PPK-positive and, when grown in MRS broth under anaerobic conditions, was shown to exhibit a small rod morphology (Fig. 2). Cells grow in anaerobic and aerobic conditions. The colony diameter, when grown on MRS agar, ranges from 1.0 to 2.5 mm. Furthermore, colonies grown on the surface of an MRS agar plate are white and circular, while the embedded colonies are white and elliptical. Moreover, after 48h the cells incubated in aerobic and

ACCEPTED MANUSCRIPT

anaerobic conditions are able to grow from 25 to 42°C, but do not grow at 20°C or 45°C (Table 1). Moreover, the cells grow at pH 5-6 (Table 1).

Fermentation profiles of *B. criceti* LMG 30188 revealed that it is able to ferment a wide range of simple and complex carbohydrates, including cellobiose, galacto-oligosaccharides, glucose, glycogen, maltodextrin, maltose and pullulan (i.e., reaching a final OD_{600} of ≥ 0.51). In contrast, *B. criceti* LMG 30188 displays little if any growth on arabinogalactan, fructo-oligosaccharides, galactose, inulin, lactose, maltotriose, mannitol, mannose, melibiose, N-acetyl-D-glucosamine, N-acetyl-D-galactosamine, raffinose, rhamnose, ribose, sorbitol, starch, sucrose, trehalose, turanose and xylose. (i.e., final OD_{600} is ≤ 0.5) (Table 2). Furthermore, the cells are not able to grow on arabinose, cellobiose, fructose and fucose. Positive enzymatic activities were observed for arginine dihydrolase, alkaline phosphatase, arginine arylamidase, glutamyl glutamic acid and serine arylamidase. Only a weak reaction was detected for leucyl-glycine arylamidase (Table 2). Moreover, α-galactosidase, β-galactosidase, α-glucosidase, β-glucosidase, α-arabinosidase, β-glucuronidase, N-acetyl-β-glucosaminidase, glutamic acid decarboxylase and pyroglutamic acid arylamidase activities were not detected (Table 2).

The peptidoglycan type is L-Lys-L-Ala_2-L-Ser. The major fatty acids identified are palmitic, myristic, oleic acid and stearic acid-dimetylacetal (Table S1).

Phylogenetic analysis of *B. criceti* LMG 30188 shows that this strain is highly related to *B. choerinum* LMG 10510 and *B. italicum* LMG 30187.

The type strain Ham19E (=LMG 30188^T=CCUG 70962^T) was isolated from the stool sample of a European hamster (*Cricetus cricetus*). The DNA G+C content is 62.53 %. Digital Protologue Taxonumber: TA00243.

Description of *Bifidobacterium anseris* sp. nov.

ACCEPTED MANUSCRIPT

Bifidobacterium anseris (an.se.ris, N.L. gen. n. *anseris* of *anser*; common scientific name of domestic goose (*Anser domesticus*)).

Cells are Gram-positive, non-motile, non-sporulating and F6PPK-positive and, when grown in MRS broth under anaerobic condition, was shown to exhibit a small rod morphology (Fig. 2). Cells grow under anaerobic and aerobic conditions. The colony diameter, when grown on MRS agar, ranges from 1.0 to 2.5 mm. Furthermore, colonies grown on the surface of an MRS agar plate are white and circular, while the embedded colonies are white and elliptical. Moreover, incubation under aerobic or anaerobic conditions show that this strain is able to grow from 25 to 42°C, yet incapable of growth at 20°C and 45°C (Table 1). Moreover, the cells grown at pH 5-6 (Table 1).

Fermentation profiles of *B. anseris* LMG 30189 revealed that it is able to ferment a wide range of simple and complex carbohydrates, including galacto-oligosaccharides, lactose, glucose, glycogen, maltodextrin, maltose, melibiose, pullulan and sucrose (i.e., reaching a final OD_{600} of ≥ 0.51). In contrast, *B. anseris* LMG 30189 displays little if any growth on arabinogalactan, fructo-oligosaccharides, galactose, maltotriose, mannitol, mannose, raffinose, rhamnose, ribose, sorbitol, starch, turanose and xylose. (i.e., final OD_{600} is ≤ 0.5) (Table 2). Furthermore, the strain does not exhibit growth on arabinose, cellobiose, fructose, fucose, inulin, N-acetyl-D-glucosamine, N-acetyl-D-galactosamine and trehalose. Positive enzymatic activities are observed for arginine dihydrolase, glutamic acid decarboxylase, alkaline phosphatase, arginine arylamidase, leucyl-glycine arylamidase and glutamyl glutamic acid (Table 2). Moreover, α-galactosidase, β-galactosidase, α-glucosidase, β-glucosidase, α-arabinosidase, β-glucuronidase, N-acetyl-β-glucosaminidase pyroglutamic acid arylamidase and serine arylamidase activity were not detected (Table 2).

The peptidoglycan type is L-Lys-L-Ser-L-Ala. The major fatty acids identified are palmitic, myristic, oleic acid and stearic acid-dimetylacetal (Table S1).

ACCEPTED MANUSCRIPT

Phylogenetic analysis of *B. anseris* LMG 30189 highlights that this strain is highly related to *B. choerinum* LMG 10510 and *B. italicum* LMG 30187.

The type strain Goo31D (=LMG 30189^T=CCUG 70960^T) was isolated from the stool sample of a domestic goose (*Anser domesticus*). The DNA G+C content is 64.3 %. Digital Protologue Taxonumber: TA00244.

Description of *Bifidobacterium parmae* sp. nov.

Bifidobacterium parmae (par.ma.e, N.L. gen. n. *parmae* of Parma; name of the city where the type strain was isolated and molecularly characterized).

Cells are Gram-positive, non-motile, non-sporulating and F6PPK-positive and, when grown in MRS broth under anaerobic condition, showed a rod morphology as well as cells with a branched or Y-shape, known as bifid-morphology (Fig. 2). Cells grow in anaerobic and aerobic conditions. The diameters of each colony, when grown on MRS agar, ranges from 1.0 to 2.0 mm. Furthermore, colonies grown on the surface of an MRS agar plate are white and circular, while embedded colonies are white and elliptical. Moreover, when incubated under aerobic or anaerobic conditions the strains is capable of growth from 25 to 42°C, yet is unable to grow at 20°C and 45°C (Table 1). Moreover, the strain was shown to grow at pH 4-6 (Table 1).

Fermentation profiles of *B. parmae* LMG 30295 revealed that it is able to ferment a wide range of simple and complex carbohydrates, including cellobiose, fructo-oligosaccharides, fructose, galactose, galacto-oligosaccharides, glucose, glycogen, inulin, lactose, maltodextrin, maltose, maltotriose, mannitol, mannose, melibiose, N-acetyl-D-glucosamine, N-acetyl-D-galactosamine, raffinose, rhamnose, ribose, sorbitol, starch, sucrose, trehalose, turanose and xylose (i.e., reaching a final OD_{600} of ≥ 0.51). In contrast, *B. parmae* LMG 30295 displays little if any growth on arabinose, arabinogalactan, fucose and pullulan (i.e., final OD_{600} is ≤ 0.5) (Table 2). Positive enzymatic activities are observed for α-arabinosidase,

ACCEPTED MANUSCRIPT

glutamic acid decarboxylase, alkaline phosphatase, arginine arylamidase, leucyl-glycine arylamidase, pyroglutamic acid arylamidase and glutamyl glutamic acid. In contrast, weak activities were observed for β-galactosidase, β-glucuronidase and serine arylamidase (Table 2). Moreover, arginine dihydrolase, α-galactosidase, α-glucosidase, β-glucosidase, and N-acetyl-β-glucosaminidase activities were not detected (Table 2).

The peptidoglycan type is L-Lys-D-Asp. The major fatty acids identified are palmitic, myristic, oleic acid and stearic acid-dimetylacetal (Table S1).

Phylogenetic analysis of *B. parmae* LMG 30295 highlights that this strain is highly related to *B. stellenboschense* DSM 23968.

The type strain Uis4E (=LMG 30295^T=CCUG 70964^T) was isolated from the stool sample of a pygmy marmoset (*Callithrix pygmaea*). The DNA G+C content is 65.81 %. Digital Protologue Taxonumber: TA00242.

Description of *Bifidobacterium margollesii* sp. nov.

Bifidobacterium margollesii (mar.gol.les'i.i N.L. gen. n. *margollesii* of Margolles; named after Professor Abelardo Margolles, in recognition of his extensive contributions to our current knowledge on bifidobacterial biology).

Cells are Gram-positive, non-motile, non-sporulating and F6PPK-positive and, when grown in MRS broth under anaerobic condition, were shown to exhibit a rod morphology as well as branched structures with a "Y" at the end, known as bifido-shaped (Fig. 2). Cells grow in anaerobic and aerobic conditions. The diameters of each colony, when grown on MRS agar, ranges from 1.0 to 2.5 mm. Furthermore, colonies grown on the surface of an MRS agar plate are white and circular, while embedded colonies are white and elliptical. Moreover, when grown under aerobic and anaerobic conditions for 48 hours the

ACCEPTED MANUSCRIPT

strain was shown to grow from 30 to 42°C, yet was not able to grow at 20°C, 25°C or 45°C (Table 1). Moreover, the cells grown at pH 4-6 (Table 1).

Fermentation profiles of *B. margollesii* LMG 30296 revealed that it is able to ferment a wide range of simple and complex carbohydrates, including arabinogalactan, cellobiose, fructo-oligosaccharides, galactose, glucose, inulin, lactose, maltotriose, mannitol, mannose, melibiose, N-acetyl-D-glucosamine, N-acetyl-D-galactosamine, ribose, sorbitol, starch, trehalose, turanose and xylose (i.e., reaching a final OD_{600} of ≥ 0.51). In contrast, *B. margollesii* LMG 30296 displays little if any growth on arabinose, fructose, fucose, galacto-oligosaccharides, glycogen, maltodextrin, maltose, pullulan, raffinose, rhamnose and sucrose (i.e., final OD_{600} is ≤ 0.5) (Table 2). Positive enzymatic activities were observed for arginine dihydrolase, alkaline phosphatase, arginine arylamidase and glutamyl glutamic acid (Table 2). Moreover, α-galactosidase, β-galactosidase, α-glucosidase, β-glucosidase, α-arabinosidase, β-glucuronidase, N-acetyl-β-glucosaminidase, glutamic acid decarboxylase, leucyl-glycine arylamidase, pyroglutamic acid arylamidase and serine arylamidase activities were not detected (Table 2).

The peptidoglycan type is L-Lys-L-Ala-L-Ser. The major fatty acids identified are palmitic, myristic, oleic acid and stearic acid-dimetylacetal (Table S1).

Phylogenetic analysis of *B. margollesii* LMG 30296 highlights that this strain is highly related to *B. hapali* DSM 100202, *B. tissieri* DSM100201 and *B. vansinderenii* LMG 30126.

The type strain Uis1B (=LMG 30296^T=CCUG 70959^T) was isolated from the stool sample of a Pygmy marmoset (*Callithrix pygmaea*). The DNA G+C content is 61.91 %. Digital Protologue Taxonumber: TA00241.

Description of *Bifidobacterium imperatoris* sp. nov.

Bifidobacterium imperatoris (i.mpe.ra.to.ris N.L. gen. n *imperatoris* of imperator; common scientific name of *Saguinus imperator*).

ACCEPTED MANUSCRIPT

Cells are Gram-positive, non-motile, non-sporulating and F6PPK-positive and, when grown in MRS broth under anaerobic conditions, exhibits a rod-shape morphology as well as a branched structure with a "Y" at the end, known as bifido-shaped (Fig. 2). Cells grow under anaerobic and aerobic conditions. The diameters of each colony, when grown on MRS agar, ranges from 1.0 to 3 mm. Furthermore, colonies grown on the surface of an MRS agar plate are white and circular, while embedded colonies are white and elliptical. Moreover, the strain when incubated under aerobic conditions for 48 hours was shown to grow from 30 to 42°C, while under anaerobic conditions it was shown to exhibit growth at a temperature range from 25°C to 42°C, yet are unable to grow at 20°C and 45°C (Table 1). Moreover, the cells exhibited growth at pH 5-6 (Table 1).

Fermentation profiles of *B. imperatoris* LMG 30297 revealed that it is able to ferment a wide range of simple and complex carbohydrates, including arabinose, fructo-oligosaccharides, galactose, glucose, glycogen lactose, galacto-oligosaccharides, glycogen, maltodextrin, maltose, maltotriose, mannitol, raffinose, ribose, sucrose, turanose and xylose (i.e., reaching a final OD_{600} of ≥ 0.51). In contrast, *B. imperatoris* LMG 30297 displays little if any growth on arabinogalactan, cellobiose, fructose, fucose, inulin, mannose, N-acetyl-D-glucosamine, N-acetyl-D-galactosamine, pullulan, rhamnose, sorbitol, starch and trehalose (i.e., final OD_{600} is ≤ 0.5) (Table 2). Positive enzymatic activities are observed for arginine dihydrolase, glutamic acid decarboxylase, alkaline phosphatase, arginine arylamidase, leucyl-glycine arylamidase, and glutamyl glutamic acid. In contrast, little activity was observed for α-arabinosidase and serine arylamidase (Table 2). Moreover, α-galactosidase, β-galactosidase, α-glucosidase, β-glucosidase, β-glucuronidase, N-acetyl-β-glucosaminidase and pyroglutamic acid arylamidase activities were not detected (Table 2).

The peptidoglycan type is L-Lys-Gly. The major fatty acids identified are palmitic, myristic, oleic acid and stearic acid-dimetylacetal (Table S1).

ACCEPTED MANUSCRIPT

Phylogenetic analysis of *B. imperatoris* LMG 30297 highlights that this strain is highly related to *B. saguini* DSM 23967.

The type strain Tam1G (=LMG 30297^T=CCUG 70961^T) was isolated from the stool sample of an emperor tamarin (*Saguinus imperator*). The DNA G+C content is 56.13 %. Digital Protologue Taxonumber: TA00245.

Conclusions

A combination of phylogenetic, genomic and phenotypic analyses allowed to differentiate strains Goo31D, Ham19E, Rab10A, Tam1G, Uis4E and Uis1B from each of the type strains of the so far recognized *Bifidobacterium* species. In this context, the combined phylogenetic analyses involving 16S rRNA gene, ITS, selected housekeeping genes and core genome sequences reveal that the isolated strains exhibit close phylogenetic relatedness with *Bifidobacterium choerinum* LMG 10510, *Bifidobacterium hapali* DSM 100202, *Bifidobacterium saguini* DSM 23967 and *Bifidobacterium stellenboschense* DSM 23968. Recently, three putative novel species, i.e., *Bifidobacterium aerophilum*, *Bifidobacterium avesanii* and *Bifidobacterium ramosum* have been proposed by Michelini et al. [12]. However, we have decided not to include these taxa in our analyses because they have not been validated by the International Committee on Systematic Bacteriology. Nevertheless, comparison between 16S partial gene sequences of these putative novel species showed identity values below 95.6 % when compared with Goo31D, Ham19E, Rab10A, Tam1G, Uis4E and Uis1B 16S sequences, clearly supporting that these letter strains are belonging to different bifidobacterial species.

The performed ANI analyses that are aimed at assessing genetic similarity levels between the six novel taxa showed strains Goo31D, Ham19E, Rab10A, Uis4E and Uis1B do not belong to any other identified bifidobacterial species sequenced so far, and are thus proposed as *Bifidobacterium anseris* sp. nov. (Goo31D = LMG 30189^T = CCUG 70960^T), *Bifidobacterium criceti* sp. nov. (Ham19E = LMG 30188^T

ACCEPTED MANUSCRIPT

= CCUG 70962^T), *Bifidobacterium italicum* sp. nov. (Rab10A = LMG 30187^T = CCUG 70963^T), *Bifidobacterium margollesii* sp. nov. (Uis1B = LMG 30296^T = CCUG 70959^T) and *Bifidobacterium parmae* sp. nov. (Uis4E = LMG 30295^T = CCUG 70964^T), respectively. Furthermore, Tam1G displays ANI percentage with *B. saguini* DSM 23967 similar to the threshold value used for species recognition [31]. Nevertheless, GGDC analysis highlights an estimated DDH value well below 70 %, so we propose to assign the name *Bifidobacterium imperatoris* sp. nov. (Tam1G = LMG 30297^T = CCUG 70961^T) to this isolate. In addition, in order to increase the number of bifidobacterial isolates and to provide a complete overview of the bifidobacterial population in other mammalian species, further isolation attempts are ongoing based on metagenomics analyses performed.

Acknowledgments

This work was funded by the EU Joint Programming Initiative – A Healthy Diet for a Healthy Life (JPI HDHL, http://www.healthydietforhealthylife.eu/) to DvS (in conjunction with Science Fondation Ireland [SFI], Grant number 15/JP-HDHL/3280) and to MV (in conjunction with MIUR, Italy). We thank GenProbio srl for financial support of the Laboratory of Probiogenomics. Part of this research is conducted using the High Performance Computing (HPC) facility of the University of Parma. LM is supported by Fondazione Cariparma, Parma, Italy. DvS is a member of The APC Microbiome Institute funded by Science Foundation Ireland (SFI), through the Irish Government's National Development Plan (Grant number SFI/12/RC/2273). The authors declare that they have no competing interests.

ACCEPTED MANUSCRIPT

References

[1] M. Ventura, C. Canchaya, A. Tauch, G. Chandra, G.F. Fitzgerald, K.F. Chater, D. van Sinderen, Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum, Microbiology and molecular biology reviews : MMBR, 71 (2007) 495-548.
[2] M. Ventura, F. Turroni, G.A. Lugli, D. van Sinderen, Bifidobacteria and humans: our special friends, from ecological to genomics perspectives, J Sci Food Agr, 94 (2014) 163-168.
[3] D.G. Bergey, W. WB, Parte AC: The Actinobacteria, in: Bergey's Manual of Systematic Bacteriology, 2012.
[4] F. Turroni, C. Milani, S. Duranti, C. Ferrario, G.A. Lugli, L. Mancabelli, D. van Sinderen, M. Ventura, Bifidobacteria and the infant gut: an example of co-evolution and natural selection, Cellular and molecular life sciences : CMLS, (2017).
[5] J.L. Round, S.K. Mazmanian, The gut microbiota shapes intestinal immune responses during health and disease, Nature reviews. Immunology, 9 (2009) 313-323.
[6] F. Turroni, V. Taverniti, P. Ruas-Madiedo, S. Duranti, S. Guglielmetti, G.A. Lugli, L. Gioiosa, P. Palanza, A. Margolles, D. van Sinderen, M. Ventura, Bifidobacterium bifidum PRL2010 modulates the host innate immune response, Applied and environmental microbiology, 80 (2014) 730-740.
[7] E. Avershina, K. Lundgard, M. Sekelja, C. Dotterud, O. Storro, T. Oien, R. Johnsen, K. Rudi, Transition from infant- to adult-like gut microbiota, Environmental microbiology, 18 (2016) 2226-2236.
[8] V. Bunesova, E. Vlkova, V. Rada, J. Killer, S. Musilova, Bifidobacteria from the gastrointestinal tract of animals: differences and similarities, Beneficial microbes, 5 (2014) 377-388.
[9] C. Milani, L. Mancabelli, G.A. Lugli, S. Duranti, F. Turroni, C. Ferrario, M. Mangifesta, A. Viappiani, P. Ferretti, V. Gorfer, A. Tett, N. Segata, D. van Sinderen, M. Ventura, Exploring Vertical Transmission of Bifidobacteria from Mother to Child, Applied and environmental microbiology, 81 (2015) 7078-7087.
[10] F. Turroni, E. Foroni, P. Pizzetti, V. Giubellini, A. Ribbera, P. Merusi, P. Cagnasso, B. Bizzarri, G.L. de'Angelis, F. Shanahan, D. van Sinderen, M. Ventura, Exploring the diversity of the bifidobacterial population in the human intestinal tract, Applied and environmental microbiology, 75 (2009) 1534-1545.
[11] S. Duranti, M. Mangifesta, G.A. Lugli, F. Turroni, R. Anzalone, C. Milani, L. Mancabelli, M.C. Ossiprandi, M. Ventura, Bifidobacterium vansinderenii sp. nov. isolated from faeces of emperor tamarin (Saguinus imperator). International Journal of Systematic and Evolutionary Microbiology., In press (2017).
[12] S. Michelini, M. Modesto, G. Filippini, C. Spiezio, C. Sandri, B. Biavati, A. Pisi, P. Mattarelli, Bifidobacterium aerophilum sp. nov., Bifidobacterium avesanii sp. nov. and Bifidobacterium ramosum sp. nov.: Three novel taxa from the faeces of cotton-top tamarin (Saguinus oedipus L.), Syst Appl Microbiol, 39 (2016) 229-236.
[13] S. Michelini, K. Oki, E. Yanokura, Y. Shimakawa, M. Modesto, P. Mattarelli, B. Biavati, K. Watanabe, Bifidobacterium myosotis sp. nov., Bifidobacterium tissieri sp. nov. and Bifidobacterium hapali sp. nov., isolated from faeces of baby common marmosets (Callithrix jacchus L.), International journal of systematic and evolutionary microbiology, 66 (2016) 255-265.
[14] J. Praet, I. Meeus, M. Cnockaert, M. Aerts, G. Smagghe, P. Vandamme, Bifidobacterium commune sp. nov. isolated from the bumble bee gut, Antonie van Leeuwenhoek, 107 (2015) 1307-1313.
[15] D. Laureys, M. Cnockaert, L. De Vuyst, P. Vandamme, Bifidobacterium aquikefiri sp. nov., isolated from water kefir, International journal of systematic and evolutionary microbiology, (2016).
[16] M. Modesto, S. Michelini, I. Stefanini, C. Sandri, C. Spiezio, A. Pisi, G. Filippini, B. Biavati, P. Mattarelli, Bifidobacterium lemurum sp nov., from faeces of the ring-tailed lemur (Lemur catta), International journal of systematic and evolutionary microbiology, 65 (2015) 1726-1734.
[17] S. Michelini, M. Modesto, A.M. Pisi, G. Filippini, C. Sandri, C. Spiezio, B. Biavati, B. Sgorbati, P. Mattarelli, Bifidobacterium. eulemuris sp. nov. isolated from the faeces of the black lemur (Eulemur macaco), International journal of systematic and evolutionary microbiology, (2016).

ACCEPTED MANUSCRIPT

[18] C. Milani, G.A. Lugli, S. Duranti, F. Turroni, F. Bottacini, M. Mangifesta, B. Sanchez, A. Viappiani, L. Mancabelli, B. Taminiau, V. Delcenserie, R. Barrangou, A. Margolles, D. van Sinderen, M. Ventura, Genomic Encyclopedia of Type Strains of the Genus Bifidobacterium, Applied and environmental microbiology, 80 (2014) 6290-6302.
[19] C. Milani, M. Mangifesta, L. Mancabelli, G.A. Lugli, K. James, S. Duranti, F. Turroni, C. Ferrario, M.C. Ossiprandi, D. van Sinderen, M. Ventura, Unveiling bifidobacterial biogeography across the mammalian branch of the tree of life, ISME J, (2017).
[20] F. Turroni, J.R. Marchesi, E. Foroni, M. Gueimonde, F. Shanahan, A. Margolles, D. van Sinderen, M. Ventura, Microbiomic analysis of the bifidobacterial population in the human distal gut, The ISME journal, 3 (2009) 745-751.
[21] M. Ventura, R. Zink, Rapid identification, differentiation, and proposed new taxonomic classification of Bifidobacterium lactis, Applied and environmental microbiology, 68 (2002) 6429-6434.
[22] C. Milani, A. Hevia, E. Foroni, S. Duranti, F. Turroni, G.A. Lugli, B. Sanchez, R. Martin, M. Gueimonde, D. van Sinderen, A. Margolles, M. Ventura, Assessing the fecal microbiota: an optimized ion torrent 16S rRNA gene-based analysis protocol, PloS one, 8 (2013) e68739.
[23] C. Milani, G.A. Lugli, F. Turroni, L. Mancabelli, S. Duranti, A. Viappiani, M. Mangifesta, N. Segata, D. van Sinderen, M. Ventura, Evaluation of bifidobacterial community composition in the human gut by means of a targeted amplicon sequencing (ITS) protocol, FEMS microbiology ecology, 90 (2014) 493-503.
[24] G.A. Lugli, C. Milani, L. Mancabelli, D. van Sinderen, M. Ventura, MEGAnnotator: a user-friendly pipeline for microbial genomes assembly and annotation, FEMS microbiology letters, 363 (2016).
[25] B. Chevreux, T. Pfisterer, B. Drescher, A.J. Driesel, W.E. Muller, T. Wetter, S. Suhai, Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs, Genome research, 14 (2004) 1147-1159.
[26] D. Hyatt, G.L. Chen, P.F. Locascio, M.L. Land, F.W. Larimer, L.J. Hauser, Prodigal: prokaryotic gene recognition and translation initiation site identification, Bmc Bioinformatics, 11 (2010) 119.
[27] Y. Zhao, J. Wu, J. Yang, S. Sun, J. Xiao, J. Yu, PGAP: pan-genomes analysis pipeline, Bioinformatics, 28 (2012) 416-418.
[28] A.J. Enright, S. Van Dongen, C.A. Ouzounis, An efficient algorithm for large-scale detection of protein families, Nucleic acids research, 30 (2002) 1575-1584.
[29] K. Katoh, K. Misawa, K. Kuma, T. Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, Nucleic acids research, 30 (2002) 3059-3066.
[30] M.A. Larkin, G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson, D.G. Higgins, Clustal W and Clustal X version 2.0, Bioinformatics, 23 (2007) 2947-2948.
[31] M. Richter, R. Rossello-Mora, Shifting the genomic gold standard for the prokaryotic species definition, Proceedings of the National Academy of Sciences of the United States of America, 106 (2009) 19126-19131.
[32] J.J. Campanella, L. Bitincka, J. Smalley, MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences, Bmc Bioinformatics, 4 (2003) 29.
[33] J.P. Meier-Kolthoff, A.F. Auch, H.P. Klenk, M. Goker, Genome sequence-based species delimitation with confidence intervals and improved distance functions, BMC bioinformatics, 14 (2013) 60.
[34] P. Schumann, Peptidoglycan Structure, Method Microbiol, 38 (2011) 101-129.
[35] L.T. Miller, Single derivatization method for routine analysis of bacterial whole-cell fatty acid methyl esters, including hydroxy acids, Journal of clinical microbiology, 16 (1982) 584-586.
[36] L.D. Kuykendall, M.A. Roy, J.J. Oneill, T.E. Devine, Fatty-Acids, Antibiotic-Resistance, and Deoxyribonucleic-Acid Homology Groups of Bradyrhizobium-Japonicum, International journal of systematic bacteriology, 38 (1988) 358-361.
[37] M. Ventura, C. Canchaya, A. Del Casale, F. Dellaglio, E. Neviani, G.F. Fitzgerald, D. van Sinderen, Analysis of bifidobacterial evolution using a multilocus approach, International journal of systematic and evolutionary microbiology, 56 (2006) 2783-2792.

ACCEPTED MANUSCRIPT

[38] G.A. Lugli, C. Milani, F. Turroni, S. Duranti, C. Ferrario, A. Viappiani, L. Mancabelli, M. Mangifesta, B. Taminiau, V. Delcenserie, D. van Sinderen, M. Ventura, Investigation of the evolutionary development of the genus Bifidobacterium by comparative genomics, Applied and environmental microbiology, 80 (2014) 6383-6394.
[39] K.T. Konstantinidis, J.M. Tiedje, Genomic insights that advance the species definition for prokaryotes, Proceedings of the National Academy of Sciences of the United States of America, 102 (2005) 2567-2572.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Figure Legends

Figure 1. Phylogenetic tree of the *Bifidobacterium* genus based on the concatenation of 233 core amino acid sequence genes of Goo31D, Ham19E, Rab10A, Tam1G, Uis4E and Uis1B and members of the *Bifidobacterium* genus. The core genes-based tree shows the subdivision of the seven phylogenetic groups of the *Bifidobacterium* genus represented with different colors. The phylogenetic tree was built by the neighbor-joining method with corresponding sequences of *Scardovia inopinata* JCM 12537 being employed as outgroup. Bootstrap percentages above 50 are shown at node points, based on 1000 replicates of the phylogenetic tree.

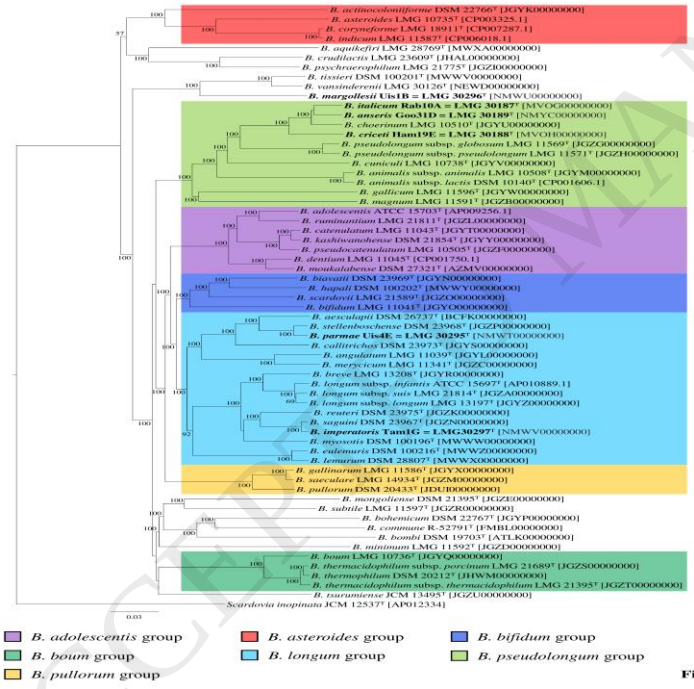


Figure 1

ACCEPTED MANUSCRIPT

Figure 2. Rab10A, Ham19E, Goo31D, Uis4E, Uis1B and Tam1G cellular morphologies as determined by the use of phase-contrast microscopy. Bar, 10 µm.

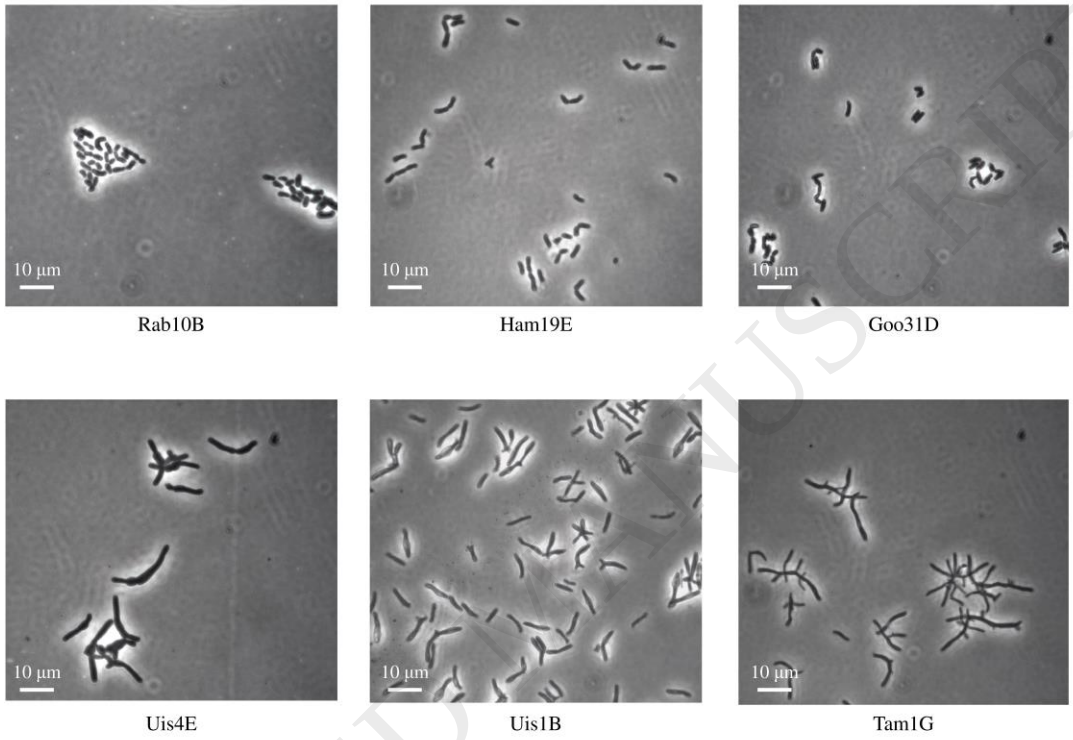


Figure 2

ACCEPTED MANUSCRIPT

Table 1. Phenotypic information for strains representing the putative novel species and their closest related species.

	Tam 1G		*B. saguini* DSM 23967		Uis 4E		Uis 1B		*B. hapali* DSM 100202		*B. tissieri* DSM 100201		*B. stellenboschense* DSM 23968		Rab10A		Ham19E		Goo31D		*B. choerinum* LMG 10510	
Temperature	**Ae**	**An**	**Ae**	**An**	**Ae**	**An**	**Ae**	**An**	**Ae**	**An**	**Ae**	**An**	**Ae**	**An**	**Ae**	**An**	**Ae**	**An**	**Ae**	**An**	**Ae**	**An**
20°C	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
25°C	-	+++	-	+	++	++	-	-	-	-	-	-	-	+	+++	+++	++	++	+	++	-	-
30°C	+++	+++	+	+	+++	+++	+	+++	+++	+++	+++	+++	+	+	+++	+++	++	++	++	++	++	++
35°C	++	+++	+++	+++	+++	+++	+++	+++	+++	++	+++	++	++++	++++	+++	+++	+++	+++	++	+++	++	+++
37°C	++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	++++	++++	+++	+++	+++	+++	++	+++	+++	++++
40°C	+++	+++	+++	+++	+++	+++	+++	+++	++	+++	+++	+++	++++	++++	+++	+++	+++	+++	++	++	++	+++
42°C	+++	+++	+++	+++	+++	+++	+	+++	+	+++	+++	+++	+	+	+++	+++	++	++	+	+	++	+++
45°C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
pH values	**24h**	**48h**	**24h**	**48h**	**24h**	**48h**	**24h**	**48h**	**24h**	**48h**	**24h**	**48h**	**24h**	**48h**	**24h**	**48h**	**24h**	**48h**	**24h**	**48h**	**24h**	**48h**
3.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+
4	-	-	-	-	+	+	++	++	+	+	-	-	+	+	+	+	+	+	+	+	+	+
4.5	-	-	-	-	++	++	+++	+++	++	++	+	++	++	++	+	+	+	+	+	+	+	+
5	+	++	+	+	++	+++	+++	+++	++	++	+++	+++	+++	+++	+	++	+	++	+	++	++	++
5.5	++	+++	+++	+++	+++	+++	+++	+++	+	++	+++	+++	+++	+++	++	+++	++	++	++	++	++	++
6	+++	+++	+++	+++	+++	+++	+++	+++	++	++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	++	+++
.																						

Ae aerobic; An anaerobic.

Symbols: -, no growth; + poor growth as measured by 0.5 ≤ OD600; ++ growth as measured by 0.51 ≤ OD600 ≤ 0.8; +++ good growth as measured by 0.81 ≤ OD600 ≤ 1.9; ++++ very good growth as measured by OD600 ≥ 1.91

ACCEPTED MANUSCRIPT

Table 2. Fermentation profiles and enzymatic activities of novel bifidobacterial species and their closest related species.

	Tam 1G		*B. saguini* DSM 23967		Uis 4E		Uis 1B		*B. hapali* DSM 100202		*B. tissieri* DSM 100201		*B. stellenboschense* DSM 23968		Rab10A		Ham19E		Goo31D		*B. choerinum* LMG 10510	
Fermentation	**24**	**48**	**24**	**48**	**24**	**48**	**24**	**48**	**24**	**48**	**24**	**48**	**24**	**48**	**24**	**48**	**24**	**48**	**24**	**48**	**24**	**48**
Arabinose	+	++	+	+	+	+	-	+	+	+	-	+	+	+	-	-	-	-	-	-	-	-
Arabinogalactan	+	+	+	+	+	+	++	++	-	+	+	+	+	+	+	+	+	+	+	+	+	+
Cellobiose	+	+	+	+	++	+++	+++	+++	+	+	+	++	+	+	-	-	-	-	-	-	-	-
Fructo-oligosaccharides	++	++	++	++	++	++	+++	+++	++	+++	+	+++	+++	+++	+	+	+	+	+	+	+	+
Fructose	+	+	+++	+++	+++	+++	+	+	++	+++	+	+++	+++	+++	-	-	-	-	-	-	-	-
Fucose	+	+	+	+	+	+	+	+	+	+	-	+	+	+	-	-	-	-	-	-	-	-
Galactose	+	++	+	++	+	+++	+++	+++	+	+	+	+++	+++	+++	+	+	+	+	+	+	+	+
Galacto-oligosaccharides	+++	+++	+++	+++	+++	+++	+	+	+	+++	+	+++	+++	+++	+	+	+	++	+	++	++	++
Glucose	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	+++	++	++	++	++	++	++	++	+++
Glycogen	++	++	+++	+++	+++	+++	+	+	-	++	-	++	++	+++	+	+	+	++	+	++	++	+++
Inulin	+	+	+	+	++	++	++	++	+	+	+	++	++	++	+	+	+	+	-	-	+	+
Lactose	++	+++	++	+++	+++	+++	++	+++	++	+++	-	+++	+++	+++	+	+	+	+	++	++	++	++
Maltodextrin	++	++	++	++	+	+++	+	+	++	++	+	++	++	++	+	++	++	++	++	++	++	+++
Maltose	+++	+++	+++	+++	+++	+++	+	+	+	+	+	+++	+++	+++	+	++	+	++	++	++	+	+++
Maltotriose	+++	+++	+++	+++	++	++	++	++	+	+++	+	++	++	++	+	+	+	+	+	+	+	++
Mannitol	+	++	+	+	+	+++	+++	+++	+	+	+	+++	+++	+++	+	+	+	+	+	+	+	+
Mannose	+	+	+	+	++	++	++	+++	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Melibiose	+	++	++	+++	+	+++	-	-	+	++	+	+++	+++	+++	+	++	+	+	+	++	+	++
N-Acetyl-D-galactosamine	+	+	+	+	++	++	++	++	+	+	+	+	+	+	-	+	+	+	-	-	-	-
N-Acetyl-D-glucosamine	+	+	+	+	++	+++	+++	+++	+	+	+	++	++	++	-	+	-	+	-	-	+	+
Pullulan	+	+	+	+	+	+	+	+	-	+	-	+	+	++	+	++	+	++	+	++	+	++
Raffionose	++	++	++	+++	+++	+++	+	+	++	++	-	+++	++	++	+	+	+	+	+	+	++	++
Rhamnose	+	+	+	+	++	++	+	+	+	+	+	+	+	+	-	-	+	+	-	+	-	-
Ribose	+	++	++	++	+	++	+++	+++	+	+	+	+	+++	+++	+	+	+	+	+	+	+	+
Sorbitol	+	+	+	+	++	+++	+++	+++	+	+	+	+	+++	+++	+	+	+	+	-	+	+	+
Starch	+	+	+	+	++	++	++	++	+	+	+	+	++	++	+	+	+	+	+	+	+	+
Sucrose	++	+++	+	++	+++	+++	-	+	++	+++	+	+++	+++	+++	+	+	+	+	++	++	++	+++
Trehalose	+	+	+	+	++	+++	++	+++	+	+	+	+++	+++	+++	+	+	+	+	-	-	+	+

ACCEPTED MANUSCRIPT

Turanose	+	++	+	++	+	+++	++	++	+	++	++	++	+++	+++	+	+	+	+	+	+	+	+
Xylose	+	++	-	-	+	+++	++	++	-	-	-	-	-	-	-	-	-	+	-	+	+	+
Enzymatic activity																						
Arginine dihydrolase	+		+		-		+		+		+		-		+		+		+		w	
α - Galactosidase	-		-		-		-		-		-		w		-		-		-		-	
β - Galactosidase	-		-		w		-		-		-		w		-		-		-		-	
α - Glucosidase	-		-		-		-		-		-		w		-		-		-		-	
β - Glucosidase	-		-		-		-		-		-		-		-		-		-		-	
α - Arabinosidase	w		+		+		-		-		-		+		-		-		-		-	
β - Glucuronidase	-		-		w		-		-		-		+		-		-		-		-	
N-acetyl-β - Glucosaminidase	-		w		-		-		-		-		-		-		-		-		w	
Glutamic acid decarboxylase	+		-		+		-		-		-		+		+		-		+		+	
Alkaline phosphatase	+		+		+		+		+		+		+		+		+		+		+	
Arginine arylamidase	+		+		+		+		+		+		+		+		+		+		+	
Leucyl-glycine arylamidase	+		+		+		-		+		+		+		+		w		+		+	
Pyroglutamic acid arylamidase	-		-		+		-		+		-		+		-		-		-		-	
Glutamyl glutamic acid arylamidase	+		+		+		+		+		-		+		+		+		+		+	
Serine arylamidase	w		+		w		-		+		-		+		w		+		-		+	
24 twenty-four hours; 48 forty-eight hours.																						
Symbols: -, no growth; + poor growth as measured by $0.5 \leq OD_{600}$; ++ growth as measured by $0.51 \leq OD_{600} \leq 0.8$; +++ good growth as measured by $0.81 \leq OD_{600} \leq 1.9$; ++++ very good growth as measured by $OD_{600} \geq 1.91$.																						
Enzymatic activity: +, Positive; -, Negative; w, Weakly Positive																						

ACCEPTED MANUSCRIPT

Table 3. General Genetic features

	Rab10A	**Ham19E**	**Goo31D**	**Tam1G**	**Uis4E**	**Uis1B**
Biological origin	European rabbit	European hamster	Domestic goose	Emperor tamarin	Pygmy marmoset	Pygmy marmoset
Average Coverage	104.3	108.6	115.4	55.52	139.27	66.93
Contigs	72	44	9	62	44	80
Genome length	2,276,351	2,155,882	2,166,761	2,639,899	2,820,211	2,789,387
Average GC percentage	65.45	62.53	64.3	56.13	65.81	61.91
Predicted ORFs	1,825	1,733	1,681	2,215	2,247	2,281
tRNA	52	53	52	59	62	60
rRNA*	4	4	4	4	6	3
16S identity	98.36% *B. choerinum*	97.32% *B. choerinum*	97.65% *B. choerinum*	99% *B. saguini*	96.8% *B. stellenboschense*	96.2% *B. tissieri*
ITS identity	87.2% *B. choerinum*	68.6% *B. animalis*	84.7% *B. choerinum*	89% *B. saguini*	81.3% *B. stellenboschense*	65.7% *B. biavatii*
ANI value	91.8% *B. choerinum*	87.81% *B. choerinum*	91.47% *B. choerinum*	94.55% *B. saguini*	93.45% *B. stellenboschense*	88.04% *B. hapali*

*Predicted number of rRNA loci