Evolutionary development and co-phylogeny of primate-associated bifidobacteria

Running title: Bifidobacterial survey in primates

Gabriele Andrea Lugli¹, Giulia Alessandri², Christian Milani^{1,3}, Leonardo Mancabelli¹, Lorena Ruiz^{4,5}, Federico Fontana¹, Santiago Borragán⁶, Andrea González⁷, Francesca Turrone^{1,3}, Maria Cristina Ossiprandi², Abelardo Margolles^{4,5}, Douwe van Sinderen⁸ and Marco Ventura^{1,3}

Laboratory of Probiogenomics, Department of Chemistry, Life Sciences, and Environmental Sustainability, University of Parma, 43124 Parma, Italy¹; Department of Veterinary Medical Science, University of Parma, 43124 Parma, Italy²; Microbiome Research Hub, University of Parma, 43124 Parma, Italy³; Department of Microbiology and Biochemistry, Dairy Research Institute of Asturias, Spanish National Research Council (IPLA-CSIC), Paseo Río Linares s/n, 33300 Villaviciosa, Asturias Spain⁴; MicroHealth Group, Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), Oviedo, Asturias, Spain⁵; Parque de Cabárceno, 39690 Obregón, Cantabria, Spain⁶; Zoo de Santillana, Avda. del Zoo 2, 39330 Santillana del Mar, Cantabria, Spain⁷; APC Microbiome Institute and School of Microbiology, Bioscience Institute, National University of Ireland, T12 YT20 Cork, Ireland⁸

gabrieleandrea.lugli@unipr.it, giulia.alessandri1@studenti.unipr.it, christian.milani@unipr.it, leonardo.mancabelli@genprobio.com, lorena.ruiz@ipla.csic.es, sborragan@cantur.com, conservacion@fundacionzoodesantillana.org, francesca.turrone@unipr.it, mariacristina.ossiprandi@unipr.it, amargolles@ipla.csic.es, d.vansinderen@ucc.ie and marco.ventura@unipr.it

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

31 **Summary**

32 In recent years, bifidobacterial populations in the gut of various monkey species have been assessed
33 in several ecological surveys, unveiling a diverse, yet unexplored ecosystem harboring novel species.
34 In the current study, we investigated the species distribution of bifidobacteria present in 23 different
35 species of primates, including human samples, by means of 16S rRNA microbial profiling and ITS
36 bifidobacterial profiling. Based on the observed bifidobacterial-host co-phylogeny, we found a
37 statistically significant correlation between the *Hominidae* family and particular bifidobacterial
38 species isolated from humans, indicating phylosymbiosis between these lineages. Furthermore,
39 phylogenetic and glycobioime analyses, based on 40 bifidobacterial species isolated from primates,
40 revealed that members of the *Bifidobacterium tissieri* phylogenetic group, which are typical gut
41 inhabitants of members of the *Cebidae* family, descend from an ancient ancestor with respect to other
42 bifidobacterial taxa isolated from primates.

43

44 **Keywords:** genomics, metagenomics, microbiota, *Bifidobacterium*

45 **Introduction**

46 The mammalian gut is inhabited by a plethora of microbial species representing the gastrointestinal
47 microbial community, also known as the gut microbiota (Lozupone et al., 2012). This microbial
48 population exerts a diversity of metabolic and physiological activities that help to sustain host health,
49 including degradation of food components, protection from pathogens, promotion of host cell
50 differentiation, and stimulation/modulation of the host immune system (Yatsunenko et al., 2012). In
51 this context, members of the genus *Bifidobacterium* are important representatives of the mammalian
52 gut microbiota, together with other key players represented by members of the *Bacteroides*,
53 *Enterococcus*, *Lactobacillus*, *Streptococcus* and *Veillonella* genera (Milani et al., 2017b).
54 Bifidobacteria are also present in other animals that confer parental care to their offspring such as
55 birds and social insects (Killer et al., 2010; Bottacini et al., 2012; Ellegaard et al., 2015). In recent
56 years, the ability of this group of microorganisms to confer a range of health benefits to the host has
57 been extensively described (Lewis et al., 2015; Arboleya et al., 2016; Turrone et al., 2016; Hidalgo-
58 Cantabrana et al., 2017). In this context, members of the *Bifidobacterium adolescentis*,
59 *Bifidobacterium animalis*, *Bifidobacterium bifidum*, *Bifidobacterium breve* and *Bifidobacterium*
60 *longum* species are known for their probiotic properties and are commonly administered as living
61 microorganisms for the (re-)establishment of a correct gut microbiota (Ventura et al., 2009; Ruiz et
62 al., 2011; Egan et al., 2014; Duranti et al., 2016).

63 Since their discovery in 1899, 16 (sub)species belonging to the genus *Bifidobacterium* have been
64 isolated and classified by means of human gut sampling (Tissier, 1900). In recent years, several
65 ecological surveys aimed at assessing bifidobacterial populations have been performed in various
66 species of monkeys, uncovering an extensive unexplored ecosystem. In fact, these efforts allowed the
67 identification of 28 novel bifidobacterial species, distributed among 11 different monkey species,
68 especially in *Saguinus oedipus* and *Callithrix jacchus*, rendering the primate gut a very rich
69 environmental repository of bifidobacterial biodiversity to date (Michellini et al., 2016; Duranti et al.,
70 2017a; Lugli et al., 2018b; Modesto et al., 2018a; Modesto et al., 2018c; Modesto et al., 2018b;

71 Duranti et al., 2019; Modesto et al., 2019). Recently, a metagenomic approach based on the
 72 amplification of the internal transcribed spacer (ITS) rRNA sequence allowed a detailed analysis of
 73 bifidobacterial gut populations in the mammalian branch of the tree of life, indicating the presence of
 74 as yet undiscovered bifidobacterial taxa (Milani et al., 2017a). Furthermore, a combination of whole
 75 metagenome shotgun sequencing coupled with cultivation approaches allowed the isolation of novel
 76 bifidobacteria from fecal samples of monkeys (Lugli et al., 2019). A comparative genomic study,
 77 involving genomes of four species of bifidobacteria isolated from marmosets and four species isolated
 78 from humans, unveiled apparently species-specific genes predicted to be involved in nutrient uptake
 79 (Brown et al., 2019).

80 In recent years, the number of studies on non-human primates have increased substantially, allowing
 81 expansion of our knowledge about their behavior, ecology and social systems (Garber, 2019).
 82 Furthermore, due to the advent of modern sequencing technologies, significant research effort has
 83 been aimed at disclosing the microbiota composition of non-human primates and at investigating how
 84 this is influenced by ecological variation through time (Greene et al., 2019; Orkin et al., 2019). These
 85 studies described gut microbiomes of several non-human primate species in which Actinobacteria were
 86 outnumbered by Firmicutes and Bacteroidetes, being reminiscent of the human gut microbiota
 87 (Bornbusch et al., 2019; Gomez et al., 2019; Greene et al., 2019). In contrast, reports on the gut
 88 microbiota of members of the *Cebidae* family, highlighted a bacterial composition dominated by the
 89 genera *Streptococcus* and *Bifidobacterium* (Orkin et al., 2019). This information is consistent with
 90 the identification of many novel bifidobacterial species from the gut microbiota of members of the
 91 *Cebidae* family.

92 In the current study, we investigated the distribution of bifidobacteria present in fecal samples from
 93 23 different species of monkeys and humans, together representing five of the major evolutionary
 94 lineages of primates, by means of 16S rRNA microbial profiling and ITS bifidobacterial profiling.
 95 The correlation between primate and *Bifidobacterium* species occurrences was used to assess possible
 96 co-phylogeny between specific hosts and their associated commensals. Furthermore, the genetic

97 features of bifidobacterial species isolated from monkeys were compared to those obtained from
98 human isolates, in order to identify differences in metabolic abilities and features related to the genetic
99 adaptation of these species to the primate gut. Finally, 40 bifidobacterial species were evaluated for
100 their glycan breakdown activities by growth measurements on carbon sources selected by *in silico*
101 predictions.

102

103 **Results and Discussion**

104 **Identification of bifidobacteria in fecal samples of monkeys.** In order to explore the relative
105 bifidobacterial abundance at genus level within the gut microbiota of various monkeys, we analyzed
106 the taxonomic composition of the bacterial fecal community harbored by 38 representatives of various
107 species of primates, of which 20 had previously been sequenced as part of a larger study aimed at
108 exploring the mammalian gut microbiota biodiversity and employing the same methodology (Milani
109 et al., 2017a) (Table S1). Altogether, the collected samples were retrieved from primates kept in
110 captivity. Illumina-mediated 16S rRNA microbial profiling of the additional 18 primate samples
111 produced approximately 800 thousand sequence reads with an average of 44 thousand quality-filtered
112 reads per sample, with a similar number obtained from the re-analysis of the previously sequenced
113 data (Table S1). The obtained nucleotide sequences were grouped in Amplicon Sequence Variants
114 (ASVs) and then taxonomically classified. This analysis showed that the relative bifidobacterial
115 abundance reached 82 % in the sample of *Callithrix argentata*, though with an average of 9 % across
116 the assessed 38 primate-derived samples (Fig. 1c). Based on primate phylogeny, we grouped each
117 monkey into families, resulting in four major clusters, i.e., *Hominidae*, *Cercopithecidae*, *Cebidae* and
118 *Lemuridae* (Table 1). Interestingly, the average abundance of bifidobacteria in primate hosts at family
119 level appears to be quite different, highlighting the *Cebidae* family as the group with the highest
120 average of relative bifidobacterial abundance (17 %), followed by the *Hominidae* family (10 %),
121 while members of the *Lemuridae* and *Cercopithecidae* families displayed relative bifidobacterial
122 abundancies of just 3 % and 0.03 %, respectively (Fig. 1c). More specifically, members of the genus

123 *Callithrix* and *Leontopithecus*, together with genus *Gorilla*, showed the highest relative abundance
124 of bifidobacteria within their deduced gut microbiota (>32 %) (Fig. 1c).

125 To compare the data collected from the non-human primate gut microbiota to that from human fecal
126 samples, 19 additional samples belonging to healthy adult individuals were analyzed following the
127 same 16S rRNA microbial profiling protocols as the monkey stool samples (Milani et al., 2019)
128 (sample size estimation of 10 and 19.5 between non-human primate species and humans, based on
129 ASVs and bacterial genera abundance, respectively). Fecal samples of humans were shown to contain
130 an average relative bifidobacterial abundance of 0.2 %, which is lower compared to the examined
131 fecal samples of monkeys, with the exception of members of the *Cercopithecidae* family (0.03 %;
132 see above). Notably, based on the α -diversity representing the whole microbiota biodiversity, human
133 samples were also shown to be lower in overall bacterial richness when compared to other analyzed
134 primate species (t-test p-value < 0.01, df = 55, Cohen's d = 0.74 and effect-size r = 0.35) (Fig. 1b).

135 Moreover, β -diversity was analyzed based on the unweighted Unifrac distance matrix and represented
136 through a principal coordinate analysis (PCoA), revealing a significant clustering of two major groups
137 represented by human and non-human primates (PERMANOVA based on 999 permutations p-value
138 = 0.001, pseudo-F = 6.20, R^2 = 0.10) (Fig. 1a). Thus, these results highlight a clear difference between
139 the microbiota composition of humans and monkeys, likely representing the outcomes of host-
140 microbiota adaptation due to multiple environmental factors and changes in the lifestyle, such as diet
141 and urbanization (De Filippo et al., 2017; Mancabelli et al., 2017; McCall et al., 2019). However, one
142 should keep in mind that the analyzed samples are derived from primates that are kept in captivity,
143 representing an unnatural environment for these animals, although their diets may have allowed
144 preservation of their bacterial gut richness.

145 To investigate the bifidobacterial distribution among various investigated primate species, we
146 employed a time tree showing the evolutionary timescale of life (Kumar et al., 2017) (Fig. 1c). Based
147 on our results, approximately 45 million years ago, the differentiation of the infraorder *Simiiformes*
148 into *Platyrrhini* and *Catarrhini* gave rise to two distinct primate lineages which now appear to harbor

149 a different abundance of members of the genus *Bifidobacterium*. In this context, the *Platyrrhini*
150 branch, represented in this study by monkeys belonging to the *Cebidae* family, appear to represent an
151 optimal ecological niche for bifidobacterial colonization when compared to members of the
152 *Catarrhini* parvorder (t-test p-value < 0.05, df = 46, Cohen's d = 0.57 and effect-size r = 0.27). Despite
153 the high number of bifidobacterial species isolated from primates, larger studies of wild primates
154 highlight that bifidobacteria do not consistently discriminate primate families from a phylogenetic
155 perspective (Amato et al., 2019; Gomez et al., 2019). In this context, phylogenetically distant primates
156 were found to possess different microbiota compositions, dominated by *Prevotella*, unknown
157 *Coriobacteriaceae* and *Streptococcus* (Gomez et al., 2019). In fact, within our analyzed samples, the
158 highest abundance of members of the genus *Bifidobacterium* was retrieved in two specific genera of
159 the *Cebidae* family, i.e., *Callithrix* and *Leontopithecus*, and not among all the members of the *Cebidae*
160 family (Fig. 1c). Thus, based on reported data, bifidobacteria seems to dominate the gut of specific
161 lineages of primates instead on a whole family of these mammals.

162 The distribution of bifidobacteria among the analyzed fecal samples of primates as highlighted by
163 16S rRNA microbial profiling, reflects from which monkey species these bifidobacteria were isolated
164 (Table 1). In fact, 26 out of 29 bifidobacterial type strains isolated from monkeys were retrieved from
165 fecal samples collected from members of the *Callitrichinae* subfamily, encompassing *Saguinus*,
166 *Callimico* and *Callithrix* genera. The remaining three bifidobacterial species were also isolated from
167 monkey-associated fecal samples that contain a substantial level of bifidobacteria, and belonging to
168 members of the *Lemuridae* and *Hominidae* families (average relative abundance of bifidobacteria of
169 3.2 and 10.3 %, respectively) (Table 1 and Fig. 1c). Interestingly, none of the bifidobacterial species
170 reported in Table 1 was isolated from the *Leontopithecus* genus, i.e., the fourth genus of the
171 *Callitrichinae* subfamily, which together with the *Saguinus*, *Callimico* and *Callithrix* genera
172 represent species with high levels of bifidobacteria within their gut microbiota.

173

Distribution of bifidobacterial species among the gut microbiota of primates. To characterize the bifidobacterial population at species level, we employed bifidobacterial ITS profiling analysis, resulting in an average of 20 thousand high-quality filtered reads per sample (Table S2). The ITS profiling approach allowed us to cluster the obtained sequences in ASVs with 100 % identity cut-off, which were subsequently classified by their bifidobacterial taxonomy at (sub)species level (Milani et al., 2014). The obtained data sets were then employed to evaluate the distribution of specific bifidobacterial species across 38 samples, defined as prevalence (Fig. 2b).

As expected, human samples revealed high prevalence (>11 %) of *B. adolescentis*, *B. longum* and *B. breve* (Fig. 2b), which are bifidobacterial species typically reported to inhabit the human gut (Turrone et al., 2018), together with other mammalian species as previously reported (Milani et al., 2017a). Furthermore, monkeys phylogenetically related to humans, i.e., those belonging to the *Hominidae*, revealed a high prevalence (>11 %) of *Bifidobacterium merycicum*, followed by *Bifidobacterium pseudolongum* and *Bifidobacterium pseudocatenulatum* that were both originally isolated from the gut of humans (Table 1). In a similar fashion, *Cercopithecidae* monkeys, belonging to the *Catarrhini* parvorder as well as *Hominidae*, disclosed a relative abundance of 70 % of the identified ITS sequences belonging to members of *B. adolescentis*, *Bifidobacterium dentium* and *B. pseudolongum* species, which were also isolated from human samples (Fig. 2b). In contrast, ITS bifidobacterial profiling revealed in monkeys of the *Cebidae* and *Lemuridae* families a high prevalence (>10 %) of *Bifidobacterium scardovii* and *Bifidobacterium ramosum*, and *Bifidobacterium lemorum* and *Bifidobacterium crudilactis*, respectively (Fig. 2b). The bifidobacterial biodiversity at species level across primates was also highlighted by the PCoA based on the unweighted Unifrac distance matrix, revealing a clustering of the *Hominidae* samples between other two groups represented by humans (PERMANOVA based on 999 permutations p-value = 0.001, pseudo-F = 5.83, $R^2 = 0.18$) and by members of the *Cebidae* family (PERMANOVA based on 999 permutations p-value = 0.002, pseudo-F = 2.69, $R^2 = 0.11$) (Fig. 2a).

199 In the same fashion as was observed for the 16S rRNA microbial profiling findings, the α -diversity
 200 based on the ITS bifidobacterial profiling highlighted a reduced bifidobacterial richness in the human
 201 samples when compared to that observed for other primates (t-test p-value < 0.001, df = 55, Cohen's
 202 d = 0.85 and effect-size r = 0.39) (Fig. 2c). In this context, the higher number of distinct bifidobacterial
 203 Operational Taxonomic Units (OTUs) identified in monkey's fecal samples was to a large degree
 204 represented by putative novel species of the genus *Bifidobacterium*, consisting of up to 22 unique
 205 OTUs next to ITS sequences of known type strains (Fig. 2b). In humans, the average relative
 206 abundance of putative novel bifidobacterial species was 6 %, while in monkey-derived samples the
 207 average percentage rises to 28 %. Looking at the primate clusters as based on family grouping,
 208 samples originating from *Cebidae* followed by *Hominidae* and *Cercopithecidae* elicited the highest
 209 levels of bifidobacterial biodiversity (ANOVA p-value < 0.01; LSD post-hoc p-value < 0.01, df = 24,
 210 Cohen's d = 2.56 and effect-size r = 0.78 and LSD post-hoc Test p-value < 0.01, df = 32, Cohen's d
 211 = 1.88 and effect-size r = 0.69 when data from *Cebidae* was compared to that from *Lemuridae* and
 212 humans, respectively) (Fig. 2d). These data were further supported by the average percentage of
 213 predicted novel bifidobacterial taxa of the *Cebidae* family, i.e., 46 %, supporting the high
 214 bifidobacterial biodiversity harbored by members of this particular primate family. Focusing on novel
 215 bifidobacterial OTUs, new_taxa_16 and new_taxa_35 were identified in 11 and eight monkey
 216 species, respectively (Table S3), followed by new_taxa_7 and new_taxa_8, both identified in samples
 217 retrieved from seven monkey species. Based on their occurrence, these putative novel
 218 *Bifidobacterium* species represent excellent targets for future projects aimed at isolating and
 219 exploring the bifidobacterial dark matter of these mammals.

220

221 **The primate-bifidobacteria co-phylogeny.** Phylogenetic relatedness among bifidobacterial species
 222 isolated from primates was investigated by means of a pangenome analysis involving available
 223 genome sequences of type strains retrieved from the NCBI database (Table 1). Based on this
 224 approach, we were able to determine putative orthologous genes among the 23 (sub)species of the

225 genus *Bifidobacterium* whose relative abundance was identified to be higher than 1 % in at least one
 226 primate-derived fecal sample (Fig. 2b and 3). The analysis resulted in the identification of 20,703
 227 clusters of orthologous genes (COGs), of which 506 were shared between all genomes, representing
 228 the core bifidobacterial coding sequences of the analyzed bifidobacterial type strains. After exclusion
 229 of paralogs, concatenation of the amino acid sequences of the remaining 462 core proteins was used
 230 to build the *Bifidobacterium* phylogenetic tree of type strains isolated from primates (Fig. 3).
 231 Furthermore, phylogenetic relationships between primate species were delineated through a time tree
 232 based on the evolutionary timescale of life (Kumar et al., 2017) (Fig. 1c and 3).
 233 In order to test the significance between individual host-bifidobacteria associations, we employed the
 234 statistical test ParaFit (Legendre et al., 2002). The co-phylogeny analysis revealed that *B. adolescentis*
 235 followed by *Bifidobacterium biavatii* and *B. ramosum*, were the taxa most commonly shared among
 236 primates (from 15 to 7), while *Leontopithecus chrysomelas* and *Callimico goeldii*, followed by
 237 *Saguinus oedipus*, were the monkey species with the highest number of associations among primate-
 238 associated bifidobacteria (from 9 to 8) (Fig. 3). Significant associations were identified among
 239 members of the *Hominidae* family, showing a correlation between all bifidobacterial species isolated
 240 from humans. Notably, the analyzed gut microbiota of *Gorilla gorilla*, *Homo sapiens* and *Pan*
 241 *troglodytes* revealed three to seven significant correlations with *B. adolescentis*, *B. bifidum*, *B. breve*,
 242 *B. catenulatum*, *B. dentium*, *B. longum* spp. and *B. pseudocatenulatum* (parafit based on 999
 243 permutations p-value < 0.05). In addition, *Gorilla gorilla* and *Pan troglodytes* were also found
 244 significantly correlated to *B. biavatii* (parafit based on 999 permutations p-value < 0.01), while
 245 *Saguinus imperator* was significantly correlated with the presence of *Bifidobacterium imperatoris*
 246 (parafit based on 999 permutations p-value < 0.05) (Fig. 3) (for more details see Table S4 and
 247 Experimental procedures).
 248 Altogether, these data provide robust evidence of phyllosymbiosis among the *Hominidae* family and
 249 bifidobacterial species isolated from humans, as underlined by co-phylogeny interactions between
 250 these lineages. In fact, all of the host-bifidobacteria links between the latter groups were shown to be

251 significant, while interactions between other lineages of primates and bifidobacteria highlighted a
252 more widespread distribution of non-human bifidobacteria.

253

254 ***Bifidobacterium* genus biogeography.** As performed in the co-phylogeny analysis, an additional
255 pangenome analysis was performed employing each bifidobacterial type strain sequenced to date,
256 including 81 genomes listed in Table S5. Based on the orthologous genes determined by this
257 approach, 255 COGs were identified to be shared among the 81 (sub)species of the genus. Following
258 the exclusion of paralogs, concatenation of 224 core protein sequences was used to build the
259 *Bifidobacterium* phylogenetic tree (Fig. 4). Among tree branches, the bifidobacterial type strains
260 isolated from primates were identified in four of the ten previously defined bifidobacterial
261 phylogenetic groups (Lugli et al., 2017; Lugli et al., 2018a; Lugli et al., 2019), i.e., *B. adolescentis*,
262 *B. bifidum*, *B. longum* and *Bifidobacterium tissieri* group. Interestingly, a small number of
263 bifidobacterial species, which had not been isolated from primates, were identified in these four
264 phylogenetic groups, i.e., *Bifidobacterium ruminantium* LMG 21811, *Bifidobacterium roussetti* DSM
265 106027, *Bifidobacterium merycicum* LMG 11341 and *B. longum* subsp. *suis* LMG 21814. Notably,
266 *Bifidobacterium gallicum* LMG 11596 was the only species isolated from primates that was not
267 included in any of the four major groups (Fig. 4).

268 Based on these results, bifidobacterial species that commonly inhabit the primate gut appear to have
269 undergone an analogous evolution that gave rise to two major branches, i.e., the *B. tissieri*
270 phylogenetic group and a second branch composed of the *B. adolescentis*, *B. bifidum* and *B. longum*
271 phylogenetic groups (Fig. 4). The first of these two major branches, defined as cluster A, includes
272 seven bifidobacterial species isolated from monkeys of the *Cebidae* family. Based on the ITS
273 bifidobacterial profiling and the co-phylogeny relationships as discussed above, four of these taxa
274 were observed to be directly correlated with hosts belonging to the *Cebidae* lineage, i.e.,
275 *Bifidobacterium callimiconis*, *Bifidobacterium catulorum*, *Bifidobacterium primatium* and
276 *Bifidobacterium vansinderenii* (Fig. 3). Instead, the second major branch was more heterogeneous in

277 terms of bifidobacterial origins, with sub-clusters based on taxa isolated from human (Clusters B and
278 D3) and non-human primates (Clusters C, D1 and D2) (Fig. 4). In this context, cluster B,
279 corresponding to the *B. adolescentis* phylogenetic group, consists of human taxa with the exception
280 of *B. moukalabense* isolated from the *Hominidae* family, while cluster C, corresponding to the *B.*
281 *bifidum* phylogenetic group, encompasses six bifidobacterial taxa retrieved from monkeys and only
282 two bifidobacterial species of human origin, i.e., *B. scardovii* and *B. bifidum*. Cluster D on the other
283 hand encompasses members of the *B. longum* phylogenetic group, including taxa that belong to both
284 human and non-human primates, producing primate sub-cluster D1 and D2, and bifidobacterial
285 species isolated from the human gut in sub-cluster D3.

286 Altogether, this phylogenetic analysis highlights that bifidobacterial strains inhabiting the gut of
287 primates are strictly related to each other. As also demonstrated by the co-phylogeny analysis, these
288 bifidobacterial strains revealed evidence of phyllosymbiosis with their primate hosts giving rise to a
289 relationship mediated by their apparent ability to colonize the gut environment of primates. In
290 contrast, members of the *B. tissieri* phylogenetic group, which are typically resident in the gut of
291 monkeys of the *Cebidae* family, seem to have descended from an ancient ancestor when compared to
292 other primate-associated bifidobacterial species. Thus, members of Cluster A, when compared to
293 other identified clusters, appear to reflect a host-species relationship that corresponds with a higher
294 relative abundance of the genus *Bifidobacterium* in the 16S rRNA microbial profiling as well as the
295 ITS bifidobacterial analysis reported in members of the *Catarrhini* parvorder.

296

297 **Insights into the carbohydrate metabolism of primate-associated bifidobacteria.** One of the key
298 genetic features of bifidobacteria, correlated with their higher ecological fitness in terms of
299 colonization of the animal gut, is their ability to utilize complex carbohydrates (Milani et al., 2015b;
300 Milani et al., 2015a). Thus, we assessed which encoded proteins possess the ability to enhance the
301 carbohydrate-harvesting abilities of *Bifidobacterium* strains isolated from the primate gut
302 environment, by means of a classification according to the Carbohydrate-Active enZyme (CAZy)

303 database (Lombard et al., 2014). The dissected proteome of the 40 bifidobacterial type strains isolated
 304 from primates (Table 1) revealed 3850 genes predicted to encode carbohydrate-active enzymes, i.e.,
 305 glycosyl hydrolases (GHs), glycosyl transferases (GTs), polysaccharide lyases (PLs), carbohydrate
 306 esterases (CEs) and carbohydrate binding modules (CBMs). Focusing on GHs, this analysis resulted
 307 in the identification of 2577 proteins predicted to possess catalytic modules involved in the
 308 degradation of carbohydrates, where GH13, GH43, GH3, GH2, GH36, GH42 and GH31
 309 outnumbered the other identified families (Fig. 5a). Interestingly, besides the absence or presence of
 310 less frequent GHs between the analyzed proteomes, each genome exhibited a varying number of
 311 predicted GH43 members, ranging from zero to 18 genes that are predicted to be involved in the
 312 degradation of complex polysaccharides, such as (arabino)xylan, (arabino)galactan and arabinan,
 313 which are main glycan constituents of the plant cell wall.

314 In order to compare genes involved in carbohydrate metabolism of the analyzed bifidobacterial
 315 species, normalization of GH counts against the total number of predicted genes was performed to
 316 generate a GH index for each type strain. The deduced GH indexes range from 1.4 % in the case of
 317 *B. gallicum* LMG 11596 to 4.81 % for *Bifidobacterium eulemuris* DSM 100216, highlighting a broad
 318 and variable GH distribution across primate-associated bifidobacteria (Table 1). The GH indexes
 319 associated with bifidobacterial species isolated from humans and those from non-human primates
 320 were shown not to significantly differ from each other. In contrast, when GH indexes were compared
 321 between members from phylogenetic clusters, significant differences were identified between cluster
 322 A and cluster B (ANOVA p-value = 0.02; LSD post-hoc p-value = 0.02, df = 11, Cohen's d = 2.43
 323 and effect-size r = 0.77), as well as between cluster A and cluster C (ANOVA p-value = 0.02; LSD
 324 post-hoc p-value = 0.005, df = 13, Cohen's d = 1.5 and effect-size r = 0.6) (Fig. 5b). As elucidated by
 325 the phylogenetic tree of the *Bifidobacterium* genus (Fig. 4), members of the *B. tissieri* group seem to
 326 have descended from an ancient ancestor. The evolutionary distance observed from this phylogenetic
 327 analysis coincides with a reduced repertoire of genes associated with the metabolism of complex
 328 carbohydrates. Despite their simpler GH repertoire, members of the *B. tissieri* group contain between

two and six GH78 family-encoding genes, which are predicted to be involved in the hydrolysis of terminal α -L-rhamnose. Besides this genetic signature, and the presence of GH28 family-encoding genes, which together with their polygalacturonase activity are also associated with the degradation of terminal α -L-rhamnose, members of this phylogenetic group revealed the absence of various accessory GHs that are present in other examined bifidobacteria (Fig. 5a). Based on these results we assume that members of the *B. tissieri* group are commensal bacteria highly abundant in the gut of the *Cebidae* family that allowed the acquisition of metabolic capabilities to degrade particular complex carbohydrates. Interestingly, among the analyzed non-human primates, some members of the *Cebidae* family, such as *Callithrix*, are gummivore (Power and Myers, 2009). Primates of the latter family, when kept in captivity are usually fed with seasonal fruits, vegetables and gum arabic, which is a biopolymer consisting in arabinogalactan provided to substitute the tree gum ingested by wild monkeys. Thus, bifidobacteria may help in the breakdown/metabolism of this complex carbon source. However, cross-feeding between different species of the genus *Bifidobacterium* or within other members of the gut microbiota could also be correlated to the identification of members of the *B. tissieri* group in the gut of the *Cebidae* family.

Glycan breakdown activities of primate-associated bifidobacteria. As mentioned above, predicted GH-encoding genes, distributed among bifidobacteria isolated from primates, suggest different carbohydrate utilization abilities between phylogenetic groups of bifidobacteria (Fig. 5). More specifically, genes encoding glycosyl hydrolases of the GH20, GH28, GH78, GH105, GH112 and GH129 families seems to be directly correlated with these significant differences between members of the *B. tissieri* group and those of the other phylogenetic groups (Fig. 5). Thus, we explored the capability of our 40 bifidobacterial type strains to utilize complex carbohydrate sources as previously performed in a similar fashion on *B. callitrichos* strains by Albert et al. (Albert et al., 2018). To validate our predictions, growth abilities of each bifidobacterial type strain isolated from the gut of primates on arabinogalactan, lacto-*N*-tetraose (LNT), or pectin as unique carbon sources,

355 were evaluated by OD_{600nm} measurements for 72 hours. While LNT was selected based on the
 356 substrate activities associated with GH20 and GH112, predicted to encode lacto-N-biosidase
 357 activities releasing lacto-*N*-biose and lactose from LNT, arabinogalactan and pectin were selected to
 358 evaluate the enzymatic activity of members of the GH28 and GH78 families, predicted to encode α -
 359 L-rhamnosidase activities releasing α -L-rhamnose substitutions from complex polysaccharides such
 360 as arabinogalactan and pectin. Notably, the listed GH families are predicted to degrade type I
 361 arabinogalactan, since type II arabinogalactan possesses a different molecular structure, which is
 362 based on a distinct backbone and branch residues.

363 As suggested by the *in silico* glycomiome analysis, strains highlighted that members of the *B. tissieri*
 364 phylogenetic group displayed higher growth performances when cultivated on arabinogalactan and
 365 pectin as unique carbon sources in comparison to members of the *B. adolescentis*, *B. bifidum* and *B.*
 366 *longum* groups (Fig. 6). Statistical analyses revealed that the breakdown capability of members of the
 367 *B. tissieri* group, based on pectin, was significant at 24h (ANOVA p-value < 0.02, LSD post-hoc p-
 368 value < 0.05) and 48h (ANOVA p-value = 0.06, LSD post-hoc p-value < 0.05), with growth
 369 performances substantially higher at 72h, reflecting enzymatic activity correlated to the presence in
 370 their genomes of genes encoding members of the GH28 and GH78 families (Fig. 6). In a similar
 371 fashion, members of the *B. tissieri* group were also shown to grow better on media containing
 372 arabinogalactan as the only carbon source when compared to other primate-associated bifidobacterial
 373 species. However, in this latter case the data was not supported by any statistical significance (Fig.
 374 6). In contrast, the growth performances of members of the *B. tissieri* group on LNT was shown to
 375 be lower than that observed for members of the *B. bifidum* (t-test p-value > 0.05 at 24h, 48h and 72h)
 376 and *B. longum* groups (t-test p-value < 0.05, df = 22, Cohen's d = 0.93 and effect-size r = 0.42 at 24h,
 377 t-test p-value < 0.05, df = 22, Cohen's d = 1.07 and effect-size r = 0.47 at 48h and t-test p-value <
 378 0.05, df = 22, Cohen's d = 1.1 and effect-size r = 0.48 at 72h) (Fig. 6). As expected, the ability to
 379 utilize human milk oligosaccharides (HMOs) such as LNT was previously observed in bifidobacterial
 380 species like *B. bifidum* PRL2010 and *B. breve* UCC2003 (Egan et al., 2014; James et al., 2016), as

well as in *B. longum* subsp. *infantis*, regarded as the archetypical HMO bifidobacterial utilizers (Sela, 2011; Garrido et al., 2015). In addition, our results showed that members of the *B. tissieri* and *B. adolescentis* groups have little if any ability to metabolize LNT (Fig. 6). Thus, based on our *in silico* and *in vitro* analyses of the glycobiome of bifidobacteria isolated from primates, we argue that arabinogalactan and pectin represent relevant growth substrates for the identification and subsequent cultivation of novel bifidobacterial species belonging to the *B. tissieri* group harbored by monkeys of the *Cebidae* family.

Conclusions

Various ecological analyses of bifidobacterial populations have been performed in monkeys, allowing the discovery of a higher number of novel species when compared to the total number of distinct bifidobacterial species isolated from humans (Duranti et al., 2017a; Lugli et al., 2018b; Modesto et al., 2018a; Modesto et al., 2018c; Modesto et al., 2018b; Duranti et al., 2019; Modesto et al., 2019). In the current study, bifidobacterial profiling of 57 fecal samples from humans and monkeys provided an overall view of bifidobacterial biodiversity present among primates. Interestingly, ITS bifidobacterial profiling clearly suggests the presence of a substantial number of as yet undiscovered bifidobacterial species in the gastrointestinal tract of monkeys, especially in members of the *Cebidae* family. This indeed points to monkeys as an important reservoir of bifidobacterial dark matter, an observation that underscores previous findings (Milani et al., 2017a). Samples collected from members of the *Leontopithecus* genus in particular disclosed a high number of putative novel species, revealing an ecological niche from which novel bifidobacterial species can realistically be isolated. Further investigations of the gut environment of other lineages of primates that were not investigated in this study will be important to complete our view of bifidobacterial biodiversity across the primate-associated tree of life.

Co-phylogeny analysis between bifidobacteria and their hosts revealed phylosymbiosis among the *Hominidae* family and bifidobacterial species isolated from humans. Furthermore, although the co-

407 phylogeny analysis did not generate additional statistically significant results, phylogenetic and
408 glycobioime investigations suggest that members of the *B. tissieri* group have evolved through the
409 improvement of their ability to colonize the gut of non-human primates belonging to the *Cebidae*
410 lineage. Furthermore, *in vitro* growth experiments were consistent with the *in silico* prediction of the
411 glycobioime of those bifidobacterial species inhabiting the gut of primates, highlighting enhanced
412 growth performances among members of the *B. tissieri* group when compared to bifidobacterial
413 species associated with the human gut. The selection of specific carbon sources, like arabinogalactan
414 and pectin, in order to isolate bifidobacterial species by means of culturomic approaches, is expected
415 to facilitate the isolation of novel species inhabiting the gut of primates. Altogether these findings are
416 expected to facilitate the discovery of bifidobacterial dark matter in primates and isolation of novel
417 bifidobacterial species in such animals.

418

419 **Experimental procedures**

420 **DNA extraction.** DNA was extracted using the QIAmp DNA Stool Mini Kit following the
421 manufacturer's instructions (Qiagen) from fecal samples collected as described in previous studies
422 (Duranti et al., 2017b; Milani et al., 2017a). DNA concentration and purity were determined
423 employing a Picodrop microtiter Spectrophotometer (Picodrop).

424

425 **Identification of bifidobacteria by 16S rRNA gene sequencing.** Partial 16S rRNA gene sequences
426 were amplified from extracted DNA using primer pair Probio_Uni/Probio_Rev, targeting the V3
427 region of the 16S rRNA gene sequence (Milani et al., 2013). Each step of the library preparation was
428 performed using HiPure Molecular Biology Grade Water (GE Healthcare, USA). In addition, a
429 negative control was sequenced in order to verify that any contamination did not occur during the
430 amplification and sequencing phases. Briefly, the negative control was processed as a normal sample
431 (see above), but HiPure Molecular Biology Grade Water was used instead of a DNA sample.
432 Furthermore, sequencing performance was validated using a synthetic mock community of eight

known organisms employing the ZymoBIOMICS HMW DNA Standard D6322 (Zymo Research, USA). Sequencing was performed using a MiSeq (Illumina, USA) instrument at the DNA sequencing facility of GenProbio srl (www.genprobio.com) according to a previously described protocol (Milani et al., 2013). Following sequencing, fastq files were processed using a custom script based on the QIIME 2 software suite (Caporaso et al., 2010). Paired-end read pairs were assembled to reconstruct the complete Probio_Uni/Probio_Rev amplicons. Quality control retained sequences with a length between 140 and 400 bp, and a mean sequence quality score of >20 while sequences with homopolymers >7 bp and mismatched primers were omitted. In order to calculate downstream diversity measures (alpha and beta diversity indices, Unifrac analysis), 16S rRNA Amplicon Sequence Variants (ASVs) were defined at 100 % sequence homology using DADA2 (Callahan et al., 2016). ASVs not encompassing at least 2 sequences of the same sample were removed. All reads were classified to the lowest possible taxonomic rank using the scikit-learn naive Bayes machine-learning classifier implemented in QIIME 2 (Bolyen et al., 2019) and a reference dataset from the SILVA database release 132 (Quast et al., 2013).

Bifidobacterial ITS sequencing. The Internally Transcribed Spacer (ITS) sequences of bifidobacteria were amplified from extracted DNA using the specific primer pair ProbioBif-ITS_Fw and ProbioBif-ITS_Rev, which targets the variable region between the 16S rRNA and 23S rRNA gene sequences (Milani et al., 2014). Sequencing was performed using a MiSeq (Illumina) instrument at the DNA sequencing facility of GenProbio srl (www.genprobio.com) according to a previously described protocol (Milani et al., 2014). Following sequencing, fastq files were processed using a custom script based on the QIIME 2 software suite (Caporaso et al., 2010). Quality control retained sequences with a length between 100 and 400 bp and mean sequence quality score of >20, while sequences with homopolymers >7 bp in length and mismatched primers were removed. In order to calculate downstream diversity measures (alpha and beta diversity indices, Unifrac analysis), ITS Operational Taxonomic Units (OTUs) were defined at 100 % sequence homology using uclust

459 (Edgar, 2010). All reads were classified to the lowest possible taxonomic rank using QIIME 2 (Milani
460 et al., 2014; Bolyen et al., 2019) and a reference dataset, consisting of an updated version of the
461 bifidobacterial ITS database (Milani et al., 2014; Milani et al., 2017a).

462

463 **Co-phylogeny between primates and bifidobacteria.** All co-phylogeny analyses were performed
464 using R version 3.5.3. Significance between individual host-bifidobacterium association was
465 performed employing the statistical test ParaFit (Legendre et al., 2002). Function “parfait” was
466 employed using 999 permutations to estimate p-values among a given host-bifidobacterium
467 association with arguments “test.links = TRUE”, “seed = NULL” and “correction = none”. Then,
468 function “cophylo” was employed to rotate and optimize visualization of both phylogenetic trees
469 (bifidobacteria and primates).

470

471 **Comparative genomics.** Open reading frames (ORFs) of both reconstructed genomes were predicted
472 with Prodigal (Hyatt et al., 2010) and annotated by means of the software MEGAnnotator (Lugli et
473 al., 2016). Two pan-genome calculations were performed using the pan-genome analysis pipeline
474 PGAP (Zhao et al., 2012), including ORFs of 81 *Bifidobacterium* genomes collected from the NCBI
475 database (Table S5). Each predicted proteome of a given bifidobacterial strain was screened for
476 orthologues against the proteome of every collected genome by means of BLASTp analysis (Altschul
477 et al., 1990) (cutoff: E value of 1×10^{-10} and 50 % identity over at least 80 % of both protein
478 sequences). The resulting output was then clustered into protein families by means of MCL (graph
479 theory-based Markov clustering algorithm) (Enright et al., 2002), using the gene family (GF) method.

480

481 **Glycobiome prediction.** The prediction of genes encoding enzymes possessing structurally related
482 catalytic and carbohydrate-binding modules catalyzing hydrolysis, modification, or synthesis of
483 glycoside bounds was performed using the CAZy database (Lombard et al., 2014). Each predicted
484 proteome of a given bifidobacterial strain was screened for orthologues against the CAZy database

485 by means of HMMER v3.3 (Wheeler and Eddy, 2013) (cutoff: E value of 1×10^{-15}) and BLASTp
486 analysis (Altschul et al., 1990) (cutoff: E value of 1×10^{-30}). A preliminary screening has been
487 performed employing dbCAN2 meta server (Zhang et al., 2018), followed by a BLASTp validation
488 of the obtained results. Predicted GHs were manually evaluated in order to remove false positive from
489 both analyses. Values of GH index were attributed dividing the number GH counts of each
490 bifidobacterial type strain against the total number of its predicted genes (Table 1).

491

492 **Strains and culture conditions.** *Bifidobacterium* strains used for carbohydrate growth assays are
493 listed in Table 1. Strains were routinely grown under anaerobic conditions in De Man, Rogosa, Sharpe
494 (MRS) medium (Sharlau) supplemented with 0.05 % L-cysteine-HCl and incubated at 37°C for 24h.
495 Anaerobic conditions were obtained by using an anaerobic cabinet (Ruskin) in which the atmosphere
496 consisted of 17 % CO₂, 80 % N₂ and 2.99 % H₂.

497

498 **Carbohydrate growth assays.** Fermentation profile experiments on different carbon sources,
499 including arabinogalactan (final concentration 0.5 %), lacto-*N*-tetraose (final concentration 0.5 %) and
500 pectin (final concentration 0.5 %) were performed using 96-well microtiter plate. Specifically,
501 after an overnight growth, turbidity was measured by a spectrophotometer (Eppendorf) and
502 bifidobacterial strains were diluted in MRS without glucose (MRS w/o glu) in order to obtain a final
503 inoculum with an OD_{600nm} of 0.1 in each well containing different sugars. Growth yields were
504 monitored by optical density at 600nm using a plate reader (BioTek, Winooski, VT, USA). The latter
505 was run for 24h, 48h and 72h, and readings were preceded by 30s shaking at medium speed. Cultures
506 were grown in biologically independent triplicates, and the resulting growth data were expressed as
507 the mean of these replicates. Carbohydrates were purchased from Sigma and Carbosynth (Berkshire,
508 UK).

509

510 **Statistical Analyses.** Statistical analyses were performed with QIIME 2 and SPSS software v. 25
511 (www.ibm.com/software/it/analytics/spss/). PERMANOVA analyses were performed using 999
512 permutations to estimate p-values for differences among populations in PCoA analyses. T-tests were
513 performed to compare differential abundance of bacterial genera, the alpha diversity and differences
514 in GHs. Furthermore, glycan breakdown activities were tested by ANOVA analysis coupled with the
515 post hoc analysis LSD (least significant difference) for multiple comparison. The sample size between
516 non-human primate species and humans was evaluated by means of Statulator
517 (<http://statulator.com/SampleSize/ss2M.html>).

518

519 **Data Deposition.** Raw sequences of 16S rRNA gene and bifidobacterial ITS profiling experiments
520 are accessible through SRA study BioProject PRJNA594910.

521

522 **Acknowledgements**

523 We thank GenProbio srl for financial support of the Laboratory of Probiogenomics. Part of this
524 research is conducted using the High Performance Computing (HPC) facility of the University of
525 Parma. The authors declare that they have no competing interests. We thank Paola Mattarelli for
526 providing us with several bifidobacterial strains isolated from primates.

527

528 **Funding**

529 This work was funded by the EU Joint Programming Initiative – A Healthy Diet for a Healthy Life
530 (JPI HDHL, <http://www.healthydietforhealthylife.eu/>) to DvS (in conjunction with Science Foundation
531 Ireland [SFI], Grant number 15/JP-HDHL/3280). DvS is a member of The APC Microbiome Institute
532 funded by Science Foundation Ireland (SFI), through the Irish Government's National Development
533 Plan (Grant numbers SFI/12/RC/2273a and SFI/12/RC/2273b).

534

535 **Authors' contributions**

536 GAL processed the metagenomic data, conducted the analyses and wrote the manuscript. GA
537 performed the experiments. CM participated in the design of the study and contributed to the
538 manuscript preparation. LM and FF contributed with the metagenomic analyses. AG and SB provided
539 the samples. LR, FT and MCO participated in the design of the study. AM and DvS participated and
540 supervised the study. MV conceived the study, participated in its design and coordination and
541 contributed to the manuscript preparation. All authors have read and approved the final manuscript.

542

543 **Conflict of Interest**

544 The authors declare that they have no competing interests.

545 References

- 546 Albert, K., Rani, A., and Sela, D.A. (2018) The comparative genomics of *Bifidobacterium callitrichos* reflects
547 dietary carbohydrate utilization within the common marmoset gut. *Microb Genom* 4.
- 548 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool.
549 *J Mol Biol* 215: 403-410.
- 550 Amato, K.R., J, G.S., Song, S.J., Nute, M., Metcalf, J.L., Thompson, L.R. et al. (2019) Evolutionary trends in
551 host physiology outweigh dietary niche in structuring primate gut microbiomes. *ISME J* 13: 576-587.
- 552 Arbolea, S., Watkins, C., Stanton, C., and Ross, R.P. (2016) Gut Bifidobacteria Populations in Human Health
553 and Aging. *Front Microbiol* 7: 1204.
- 554 Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A. et al. (2019)
555 Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*
556 37: 852-857.
- 557 Bornbusch, S.L., Greene, L.K., McKenney, E.A., Volkoff, S.J., Midani, F.S., Joseph, G. et al. (2019) A
558 comparative study of gut microbiomes in captive nocturnal strepsirrhines. *Am J Primatol* 81: e22986.
- 559 Bottacini, F., Milani, C., Turrone, F., Sanchez, B., Foroni, E., Duranti, S. et al. (2012) *Bifidobacterium*
560 *asteroides* PRL2011 genome analysis reveals clues for colonization of the insect gut. *PLoS One* 7: e44229.
- 561 Brown, C.J., Mtui, D., Oswald, B.P., Van Leuven, J.T., Vallender, E.J., Schultz-Darken, N. et al. (2019)
562 Comparative genomics of *Bifidobacterium* species isolated from marmosets and humans. *Am J Primatol*
563 81: e983.
- 564 Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J., and Holmes, S.P. (2016) DADA2: High-
565 resolution sample inference from Illumina amplicon data. *Nat Methods* 13: 581-583.
- 566 Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K. et al. (2010) QIIME
567 allows analysis of high-throughput community sequencing data. *Nat Methods* 7: 335-336.
- 568 De Filippo, C., Di Paola, M., Ramazzotti, M., Albanese, D., Pieraccini, G., Banci, E. et al. (2017) Diet,
569 Environments, and Gut Microbiota. A Preliminary Investigation in Children Living in Rural and Urban
570 Burkina Faso and Italy. *Front Microbiol* 8: 1979.
- 571 Duranti, S., Mangifesta, M., Lugli, G.A., Turrone, F., Anzalone, R., Milani, C. et al. (2017a) *Bifidobacterium*
572 *vansinderenii* sp. nov., isolated from faeces of emperor tamarin (*Saguinus imperator*). *Int J Syst Evol*
573 *Microbiol*.
- 574 Duranti, S., Lugli, G.A., Napoli, S., Anzalone, R., Milani, C., Mancabelli, L. et al. (2019) Characterization of
575 the phylogenetic diversity of five novel species belonging to the genus *Bifidobacterium*: *Bifidobacterium*
576 *castoris* sp. nov., *Bifidobacterium callimiconis* sp. nov., *Bifidobacterium goeldii* sp. nov., *Bifidobacterium*
577 *samirii* sp. nov. and *Bifidobacterium dolichotidis* sp. nov. *Int J Syst Evol Microbiol*.
- 578 Duranti, S., Milani, C., Lugli, G.A., Mancabelli, L., Turrone, F., Ferrario, C. et al. (2016) Evaluation of genetic
579 diversity among strains of the human gut commensal *Bifidobacterium adolescentis*. *Sci Rep* 6: 23971.
- 580 Duranti, S., Lugli, G.A., Mancabelli, L., Armanini, F., Turrone, F., James, K. et al. (2017b) Maternal inheritance
581 of bifidobacterial communities and bifidophages in infants through vertical transmission. *Microbiome* 5:
582 66.
- 583 Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460-
584 2461.
- 585 Egan, M., Motherway, M.O., Kilcoyne, M., Kane, M., Joshi, L., Ventura, M., and van Sinderen, D. (2014)
586 Cross-feeding by *Bifidobacterium breve* UCC2003 during co-cultivation with *Bifidobacterium bifidum*
587 PRL2010 in a mucin-based medium. *BMC Microbiol* 14: 282.
- 588 Ellegaard, K.M., Tamarit, D., Javelind, E., Olofsson, T.C., Andersson, S.G., and Vasquez, A. (2015) Extensive
589 intra-phylogroup diversity in lactobacilli and bifidobacteria from the honeybee gut. *BMC Genomics* 16: 284.
- 590 Enright, A.J., Van Dongen, S., and Ouzounis, C.A. (2002) An efficient algorithm for large-scale detection of
591 protein families. *Nucleic Acids Res* 30: 1575-1584.
- 592 Garber, P.A. (2019) Distinguished primatologist address-moving from advocacy to activism: Changing views
593 of primate field research and conservation over the past 40 years. *Am J Primatol* 81: e23052.

594 Garrido, D., Ruiz-Moyano, S., Lemay, D.G., Sela, D.A., German, J.B., and Mills, D.A. (2015) Erratum:
595 Comparative transcriptomics reveals key differences in the response to milk oligosaccharides of infant gut-
596 associated bifidobacteria. *Sci Rep* 5: 15311.

597 Gomez, A., Sharma, A.K., Mallott, E.K., Petrzalkova, K.J., Jost Robinson, C.A., Yeoman, C.J. et al. (2019)
598 Plasticity in the Human Gut Microbiome Defies Evolutionary Constraints. *mSphere* 4.

599 Greene, L.K., Bornbusch, S.L., McKenney, E.A., Harris, R.L., Gorvetzian, S.R., Yoder, A.D., and Drea, C.M.
600 (2019) The importance of scale in comparative microbiome research: New insights from the gut and glands
601 of captive and wild lemurs. *Am J Primatol* 81: e22974.

602 Hidalgo-Cantabrana, C., Delgado, S., Ruiz, L., Ruas-Madiedo, P., Sanchez, B., and Margolles, A. (2017)
603 Bifidobacteria and Their Health-Promoting Effects. *Microbiol Spectr* 5.

604 Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal: prokaryotic
605 gene recognition and translation initiation site identification. *BMC Bioinformatics* 11: 119.

606 James, K., Motherway, M.O., Bottacini, F., and van Sinderen, D. (2016) Bifidobacterium breve UCC2003
607 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-tetraose through
608 overlapping, yet distinct pathways. *Sci Rep* 6: 38560.

609 Killer, J., Kopecký, J., Mrazek, J., Rada, V., Dubna, S., and Marounek, M. (2010) Bifidobacteria in the
610 digestive tract of bumblebees. *Anaerobe* 16: 165-170.

611 Kumar, S., Stecher, G., Suleski, M., and Hedges, S.B. (2017) TimeTree: A Resource for Timelines, Timetrees,
612 and Divergence Times. *Mol Biol Evol* 34: 1812-1819.

613 Legendre, P., Desdevises, Y., and Bazin, E. (2002) A statistical test for host-parasite coevolution. *Syst Biol*
614 51: 217-234.

615 Lewis, Z.T., Totten, S.M., Smilowitz, J.T., Popovic, M., Parker, E., Lemay, D.G. et al. (2015) Maternal
616 fucosyltransferase 2 status affects the gut bifidobacterial communities of breastfed infants. *Microbiome* 3:
617 13.

618 Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014) The carbohydrate-
619 active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42: D490-495.

620 Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., and Knight, R. (2012) Diversity, stability and
621 resilience of the human gut microbiota. *Nature* 489: 220-230.

622 Lugli, G.A., Milani, C., Mancabelli, L., van Sinderen, D., and Ventura, M. (2016) MEGAnnotator: a user-
623 friendly pipeline for microbial genomes assembly and annotation. *FEMS Microbiol Lett* 363.

624 Lugli, G.A., Milani, C., Duranti, S., Mancabelli, L., Mangifesta, M., Turrone, F. et al. (2018a) Tracking the
625 Taxonomy of the Genus Bifidobacterium Based on a Phylogenomic Approach. *Appl Environ Microbiol* 84.

626 Lugli, G.A., Milani, C., Duranti, S., Alessandri, G., Turrone, F., Mancabelli, L. et al. (2019) Isolation of novel
627 gut bifidobacteria using a combination of metagenomic and cultivation approaches. *Genome Biol* 20: 96.

628 Lugli, G.A., Mangifesta, M., Duranti, S., Anzalone, R., Milani, C., Mancabelli, L. et al. (2018b) Phylogenetic
629 classification of six novel species belonging to the genus Bifidobacterium comprising Bifidobacterium
630 anseris sp. nov., Bifidobacterium criceti sp. nov., Bifidobacterium imperatoris sp. nov., Bifidobacterium
631 italicum sp. nov., Bifidobacterium margollesii sp. nov. and Bifidobacterium parmae sp. nov. *Syst Appl*
632 *Microbiol* 41: 173-183.

633 Lugli, G.A., Milani, C., Turrone, F., Duranti, S., Mancabelli, L., Mangifesta, M. et al. (2017) Comparative
634 genomic and phylogenomic analyses of the Bifidobacteriaceae family. *BMC Genomics* 18: 568.

635 Mancabelli, L., Milani, C., Lugli, G.A., Turrone, F., Ferrario, C., van Sinderen, D., and Ventura, M. (2017)
636 Meta-analysis of the human gut microbiome from urbanized and pre-agricultural populations. *Environ*
637 *Microbiol* 19: 1379-1390.

638 McCall, L.I., Callewaert, C., Zhu, Q., Song, S.J., Bouslimani, A., Minich, J.J. et al. (2019) Home chemical and
639 microbial transitions across urbanization. *Nat Microbiol*.

640 Michelini, S., Modesto, M., Filippini, G., Spiezio, C., Sandri, C., Biavati, B. et al. (2016) Bifidobacterium
641 aerophilum sp. nov., Bifidobacterium avesanii sp. nov. and Bifidobacterium ramosum sp. nov.: Three novel
642 taxa from the faeces of cotton-top tamarin (*Saguinus oedipus* L.). *Syst Appl Microbiol* 39: 229-236.

643 Milani, C., Turrone, F., Duranti, S., Lugli, G.A., Mancabelli, L., Ferrario, C. et al. (2015a) Genomics of the
644 Genus Bifidobacterium Reveals Species-Specific Adaptation to the Glycan-Rich Gut Environment. *Appl*
645 *Environ Microbiol* 82: 980-991.

646 Milani, C., Lugli, G.A., Turrone, F., Mancabelli, L., Duranti, S., Viappiani, A. et al. (2014) Evaluation of
647 bifidobacterial community composition in the human gut by means of a targeted amplicon sequencing
648 (ITS) protocol. *FEMS Microbiol Ecol* 90: 493-503.

649 Milani, C., Mangifesta, M., Mancabelli, L., Lugli, G.A., James, K., Duranti, S. et al. (2017a) Unveiling
650 bifidobacterial biogeography across the mammalian branch of the tree of life. *ISME J* 11: 2834-2847.

651 Milani, C., Hevia, A., Foroni, E., Duranti, S., Turrone, F., Lugli, G.A. et al. (2013) Assessing the fecal
652 microbiota: an optimized ion torrent 16S rRNA gene-based analysis protocol. *PLoS One* 8: e68739.

653 Milani, C., Duranti, S., Napoli, S., Alessandri, G., Mancabelli, L., Anzalone, R. et al. (2019) Colonization of
654 the human gut by bovine bacteria present in Parmesan cheese. *Nat Commun* 10: 1286.

655 Milani, C., Duranti, S., Bottacini, F., Casey, E., Turrone, F., Mahony, J. et al. (2017b) The First Microbial
656 Colonizers of the Human Gut: Composition, Activities, and Health Implications of the Infant Gut
657 Microbiota. *Microbiol Mol Biol Rev* 81.

658 Milani, C., Lugli, G.A., Duranti, S., Turrone, F., Mancabelli, L., Ferrario, C. et al. (2015b) Bifidobacteria exhibit
659 social behavior through carbohydrate resource sharing in the gut. *Sci Rep* 5: 15782.

660 Modesto, M., Michelini, S., Oki, K., Biavati, B., Watanabe, K., and Mattarelli, P. (2018a) Bifidobacterium
661 catulorum sp. nov., a novel taxon from the faeces of the baby common marmoset (*Callithrix jacchus*). *Int J*
662 *Syst Evol Microbiol* 68: 575-581.

663 Modesto, M., Puglisi, E., Bonetti, A., Michelini, S., Spiezio, C., Sandri, C. et al. (2018b) Bifidobacterium
664 primatium sp. nov., Bifidobacterium scaligerum sp. nov., Bifidobacterium felsineum sp. nov. and
665 Bifidobacterium simiarum sp. nov.: Four novel taxa isolated from the faeces of the cotton top tamarin
666 (*Saguinus oedipus*) and the emperor tamarin (*Saguinus imperator*). *Syst Appl Microbiol* 41: 593-603.

667 Modesto, M., Watanabe, K., Arita, M., Satti, M., Oki, K., Sciavilla, P. et al. (2019) Bifidobacterium jacchi sp.
668 nov., isolated from the faeces of a baby common marmoset (*Callithrix jacchus*). *Int J Syst Evol Microbiol*
669 69: 2477-2485.

670 Modesto, M., Michelini, S., Sansosti, M.C., De Filippo, C., Cavalieri, D., Qvirist, L. et al. (2018c)
671 Bifidobacterium callitrichidarum sp. nov. from the faeces of the emperor tamarin (*Saguinus imperator*). *Int*
672 *J Syst Evol Microbiol* 68: 141-148.

673 Orkin, J.D., Campos, F.A., Myers, M.S., Cheves Hernandez, S.E., Guadamuz, A., and Melin, A.D. (2019)
674 Seasonality of the gut microbiota of free-ranging white-faced capuchins in a tropical dry forest. *ISME J* 13:
675 183-196.

676 Power, M.L., and Myers, E.W. (2009) Digestion in the common marmoset (*Callithrix jacchus*), a gummivore-
677 frugivore. *Am J Primatol* 71: 957-963.

678 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P. et al. (2013) The SILVA ribosomal RNA
679 gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41: D590-596.

680 Ruiz, L., Gueimonde, M., Coute, Y., Salminen, S., Sanchez, J.C., de los Reyes-Gavilan, C.G., and Margolles,
681 A. (2011) Evaluation of the ability of Bifidobacterium longum to metabolize human intestinal mucus. *FEMS*
682 *Microbiol Lett* 314: 125-130.

683 Sela, D.A. (2011) Bifidobacterial utilization of human milk oligosaccharides. *Int J Food Microbiol* 149: 58-
684 64.

685 Tissier, H. (1900) Recherches sur la flore intestinale des nourrissons (etat normal et pathologique). In:
686 University of Paris, Paris, France.

687 Turrone, F., Milani, C., Duranti, S., Ferrario, C., Lugli, G.A., Mancabelli, L. et al. (2018) Bifidobacteria and the
688 infant gut: an example of co-evolution and natural selection. *Cell Mol Life Sci* 75: 103-118.

689 Turrone, F., Milani, C., Duranti, S., Mancabelli, L., Mangifesta, M., Viappiani, A. et al. (2016) Deciphering
690 bifidobacterial-mediated metabolic interactions and their impact on gut microbiota by a multi-omics
691 approach. *ISME J* 10: 1656-1668.

692 Ventura, M., O'Flaherty, S., Claesson, M.J., Turrone, F., Klaenhammer, T.R., van Sinderen, D., and O'Toole,
693 P.W. (2009) Genome-scale analyses of health-promoting bacteria: probiogenomics. *Nat Rev Microbiol* 7:
694 61-71.

695 Wheeler, T.J., and Eddy, S.R. (2013) nhmmer: DNA homology search with profile HMMs. *Bioinformatics* 29:
696 2487-2489.

697 Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M. et al. (2012)
698 Human gut microbiome viewed across age and geography. *Nature* 486: 222-227.
699 Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z. et al. (2018) dbCAN2: a meta server for
700 automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 46: W95-W101.
701 Zhao, Y., Wu, J., Yang, J., Sun, S., Xiao, J., and Yu, J. (2012) PGAP: pan-genomes analysis pipeline.
702 *Bioinformatics* 28: 416-418.

703

704

705 **Figure legends**

706 **Figure 1.** 16S rRNA gene-based microbial profiling of 57 fecal samples from humans and various
707 non-human primates. Panel a shows the PCoA representation of the beta-diversity related to samples
708 which are colored based on corresponding primate lineages. Polygons delimiting each group were
709 manually generated. Panel b exhibits a Whisker plot based on Chao1 alpha-diversity of non-human
710 primate and human gut microbiota data, which results in a P value of < 0.004 between the two groups
711 (Student's t test). The y axis shows the number of observed ASVs obtained from fecal samples of
712 non-human primates (grey) or humans (purple). Dots reflect the distribution of a data set, while the
713 boxes represent 50 % of the data set, distributed between the 1st and 3rd quartiles. The median divides
714 the boxes into the interquartile range, while the X represents the mean. The lines extending vertically
715 outside the boxes show the outlier range. Panel c displays a tree based on the evolutionary timescale
716 of life including families of the analyzed primates (left), and the relative abundance of bifidobacterial
717 DNA retrieved by means of 16S rRNA gene profiling (right, green pillars). Tree branches are colored
718 based on the primate lineages: *Atelidae* (red), *Cebidae* (blue), *Cercopithecidae* (orange), *Hominidae*
719 (green) and *Lemuridae* (yellow).

720

721 **Figure 2.** ITS bifidobacterial profiling of 57 fecal samples of humans and various non-human primate
722 species. Panel a visualizes the PCoA representation of the beta-diversity related to samples colored
723 based on primate lineages. Polygons delimiting each group were manually generated. Panel b displays
724 the relative abundance of each identified species belonging to the genus *Bifidobacterium* by means
725 of ITS bifidobacterial profiling. Numbers represent the average percentage of bifidobacterial species
726 identified in a specific primate family. Only species that display at least 1 % of the total amount of
727 data in at least a primate family are included in the heat map. The average percentage of putative
728 novel bifidobacterial species are reported together, while order of *Bifidobacterium* species is reported
729 based on their abundance among primate family. Panels c and d exhibit Whisker plots based on Chao1
730 alpha-diversity of non-human primate and human gut microbiota data. The y axis shows the number

731 of observed OTUs obtained from fecal samples of non-human primates (grey) or humans (purple),
732 while family groups were colored based on primate lineages. Dots reflect the distribution of a data
733 set, while the boxes represent 50 % of the data set, distributed between the 1st and 3rd quartiles. The
734 median divides the boxes into the interquartile range, while the X represents the mean. The lines
735 extending vertically outside the boxes show the outlier range.

736

737 **Figure 3.** Tanglegram of co-phylogenetic relationships between 24 primates and 23 bifidobacterial
738 species. Tanglegram is composed by a tree based on the evolutionary timescale of families of primates
739 (left) and a proteomic tree based on the concatenation of 462 core genes identified in the pangenome
740 analysis of 23 *Bifidobacterium* type strains (right). All associations are shown in the tanglegram as
741 black and red connecting lines. The latter lines indicate significant individual co-speciation links
742 between bifidobacteria and their hosts as indicated by ParaFit ($P < 0.05$), while black lines represent
743 non-significant links. Bifidobacterial and primate trees are rooted with *Scardovia inopinata* JCM
744 12537 and *Cynocephalus volans* as outgroups, respectively. Bootstrap percentages, of the
745 bifidobacterial proteome tree, above 50 are shown at node points, based on 1,000 replicates.

746

747 **Figure 4.** Phylogenomic tree of the *Bifidobacterium* genus. The proteomic tree was based on
748 concatenation of 224 core genes identified in the pangenome analysis of 81 *Bifidobacterium* strains.
749 Phylogenetic groups are highlighted in different colors. Bifidobacterial species isolated from primates
750 are highlighted based on their isolation source, i.e., human (teal) and non-human primate (brown),
751 while clusters are reported with alphabetical letters. The tree was constructed by the neighbor-joining
752 method, and the genome sequence of *Scardovia inopinata* JCM 12537 was used as an outgroup.
753 Bootstrap percentages above 50 are shown at node points, based on 1,000 replicates.

754

755 **Figure 5.** Comparative analysis of bifidobacterial GHs. Panel a shows a heat map of the GH
756 prediction among 40 *Bifidobacterium* species isolated from primates. Panel b displays a Whisker plot

757 based on the GH index identified for each *Bifidobacterium* species. The y axis shows the number of
758 the GH index, subdivided in four major cluster highlighted with different colors based on their
759 phylogenetic groups. Dots reflect the distribution of a data set, while the boxes represent 50 % of the
760 data set, distributed between the 1st and 3rd quartiles. The median divides the boxes into the
761 interquartile range, while the X represents the mean. The lines extending vertically outside the boxes
762 show the outlier range. Statistical significance is reported by connection between groups.

763

764 **Figure 6.** Growth experiments based on MRS supplanted as unique carbon source by arabinogalactan,
765 lacto-*N*-tetraose and pectin. Panel a shows a heat map of the growth yields among 40 *Bifidobacterium*
766 species isolated from primates. Panel b displays six Whisker plots based on the growth experiments
767 of the bifidobacterial species on arabinogalactan and pectin at 24h, 48h and 72h. The y axis shows
768 the optical density values obtained, subdivided in four major cluster highlighted with different colors
769 based on their phylogenetic groups. Dots reflect the distribution of a data set, while the boxes
770 represent 50 % of the data set, distributed between the 1st and 3rd quartiles. The median divides the
771 boxes into the interquartile range, while the X represents the mean. The lines extending vertically
772 outside the boxes show the outlier range. Statistical significance is reported by connection between
773 groups.

774

775 **Supporting information**

776 **Supplementary table 1.** Quality-filtering table of 16S rRNA gene profiling datasets.

777 **Supplementary table 2.** Quality-filtering table of ITS bifidobacterial profiling datasets.

778 **Supplementary table 3.** ITS bifidobacterial profiling of primate samples.

779 **Supplementary table 4.** ParaFit significance between individual host-bifidobacterium associations.

780 **Supplementary table 5.** List of *Bifidobacterium* type strains.

Table 1. General genomic features of bifidobacteria isolated from primates.

N°	<i>Bifidobacterium</i> species	Strain	Genome Status	Genome Size	GC content	ORFs Number	rRNA loci*	tRNA Number	GHs Number	GH Index	Accession Number	Isolated from	Host species	Host Family
1	<i>Bifidobacterium adolescentis</i>	ATCC 15703	Complete	2089645	59.18	1676	5	54	56	3.4%	AP009256.1	Human	<i>Homo sapiens</i>	<i>Hominidae</i>
2	<i>Bifidobacterium aerophilum</i>	DSM 100689	-	-	-	-	-	-	-	-	-	Cotton-top tamarin	<i>Saguinus oedipus</i>	<i>Cebidae</i>
3	<i>Bifidobacterium aesculapii</i>	DSM 26737	Draft (118)	2794396	64.58	2151	6	60	57	2.6%	BCFK000000000	Common marmoset	<i>Callithrix jacchus</i>	<i>Cebidae</i>
4	<i>Bifidobacterium angulatum</i>	DSM 20098	Complete	2021974	59.43	1557	4	53	38	2.5%	AP012322.1	Human	<i>Homo sapiens</i>	<i>Hominidae</i>
5	<i>Bifidobacterium avesanii</i>	DSM 100685	Draft (46)	2684414	66.29	2046	4	56	79	3.9%	WBSN000000000	Cotton-top tamarin	<i>Saguinus oedipus</i>	<i>Cebidae</i>
6	<i>Bifidobacterium biavatii</i>	DSM 23969	Draft (56)	3252147	63.1	2557	5	61	107	4.2%	JGYN000000000	Red-handed tamarin	<i>Saguinus midas</i>	<i>Cebidae</i>
7	<i>Bifidobacterium bifidum</i>	ATCC 29521	Complete	2211039	62.66	1822	2	54	41	2.3%	AP012323.1	Human	<i>Homo sapiens</i>	<i>Hominidae</i>
8	<i>Bifidobacterium breve</i>	DSM 20213	Complete	2269415	58.89	1977	2	53	43	2.2%	AP012324.1	Human	<i>Homo sapiens</i>	<i>Hominidae</i>
9	<i>Bifidobacterium callimiconis</i>	LMG 30938	Draft (32)	2962404	62.37	2263	5	61	61	2.7%	QXGJ000000000	Goeldi's marmoset	<i>Callimico goeldii</i>	<i>Cebidae</i>
10	<i>Bifidobacterium callitrichidarum</i>	DSM 103152	Draft (58)	3121265	61.81	2526	2	79	73	2.9%	QFFM000000000	Emperor tamarin	<i>Saguinus imperator</i>	<i>Cebidae</i>
11	<i>Bifidobacterium callitrichos</i>	DSM 23973	Draft (33)	2887313	63.52	2364	3	58	81	3.4%	JGYS000000000	Common marmoset	<i>Callithrix jacchus</i>	<i>Cebidae</i>
12	<i>Bifidobacterium catenulatum</i> subsp. <i>catenulatum</i>	DSM 16992	Complete	2079525	56.2	1686	5	55	52	3.1%	AP012325.1	Human	<i>Homo sapiens</i>	<i>Hominidae</i>
13	<i>Bifidobacterium catenulatum</i> subsp. <i>kashiwanohense</i>	DSM 21854	Complete	2337234	56.28	1932	5	54	59	3.1%	AP012327.1	Human	<i>Homo sapiens</i>	<i>Hominidae</i>
14	<i>Bifidobacterium catulorum</i>	DSM 103154	Draft (78)	2611484	63.17	2043	2	55	33	1.6%	QFFN000000000	Common marmoset	<i>Callithrix jacchus</i>	<i>Cebidae</i>
15	<i>Bifidobacterium dentium</i>	JCM 1195	Complete	2635669	58.54	2167	4	56	84	3.9%	AP012326.1	Human	<i>Homo sapiens</i>	<i>Hominidae</i>
16	<i>Bifidobacterium eulemuris</i>	DSM 100216	Draft (34)	2913389	62.2	2328	2	53	115	4.9%	MWWZ000000000	Black lemur	<i>Eulemur macaco</i>	<i>Lemuridae</i>
17	<i>Bifidobacterium felsineum</i>	DSM 103139	Draft (12)	2382257	57.11	1872	2	56	61	3.3%	PEBJ000000000	Cotton-top tamarin	<i>Saguinus oedipus</i>	<i>Cebidae</i>
18	<i>Bifidobacterium gallicum</i>	LMG 11596	Draft (12)	2004594	57.61	1505	2	58	22	1.5%	JGYW000000000	Human	<i>Homo sapiens</i>	<i>Hominidae</i>
19	<i>Bifidobacterium goeldii</i>	LMG 30939	Draft (26)	2608375	56.08	2051	2	51	64	3.1%	QXGL000000000	Goeldi's marmoset	<i>Callimico goeldii</i>	<i>Cebidae</i>
20	<i>Bifidobacterium hapali</i>	DSM 100202	Draft (76)	2834308	54.5	2246	3	54	96	4.3%	MWVY000000000	Common marmoset	<i>Callithrix jacchus</i>	<i>Cebidae</i>
21	<i>Bifidobacterium imperatoris</i>	LMG 30297	Draft (62)	2639899	56.13	2203	4	59	62	2.8%	NM WV000000000	Emperor tamarin	<i>Saguinus imperator</i>	<i>Cebidae</i>
22	<i>Bifidobacterium jacchi</i>	DSM 103362	Draft (100)	2851761	62.26	2078	1	55	43	2.1%	RQSP000000000	Common marmoset	<i>Callithrix jacchus</i>	<i>Cebidae</i>
23	<i>Bifidobacterium lemumum</i>	DSM 28807	Draft (38)	2944293	62.64	2316	3	49	109	4.7%	MWWX000000000	Ring-tailed lemur	<i>Lemur catta</i>	<i>Lemuridae</i>
24	<i>Bifidobacterium longum</i> subsp. <i>infantis</i>	ATCC 15697	Complete	2832748	59.86	2580	4	77	51	2.0%	CP001095.1	Human	<i>Homo sapiens</i>	<i>Hominidae</i>
25	<i>Bifidobacterium longum</i> subsp. <i>longum</i>	JCM 1217	Complete	2385164	60.33	2001	4	73	56	2.8%	AP010888.1	Human	<i>Homo sapiens</i>	<i>Hominidae</i>
26	<i>Bifidobacterium margollesii</i>	LMG 30296	Draft (80)	2789387	61.91	2270	3	60	53	2.3%	NMWU000000000	Pygmy marmoset	<i>Callithrix pygmaea</i>	<i>Cebidae</i>

27	<i>Bifidobacterium moukalabense</i>	DSM 27321	Draft (12)	2515335	59.87	2046	4	56	85	4.2%	AZMV000000000	Western lowland gorilla	<i>Gorilla gorilla</i>	<i>Hominidae</i>
28	<i>Bifidobacterium myosotis</i>	DSM 100196	Draft (58)	2944195	62.55	2163	4	56	76	3.5%	MWWW000000000	Common marmoset	<i>Callithrix jacchus</i>	<i>Cebidae</i>
29	<i>Bifidobacterium parmae</i>	LMG 30295	Draft (44)	2820211	65.81	2243	6	62	43	1.9%	NMWT000000000	Pygmy marmoset	<i>Callithrix pygmaea</i>	<i>Cebidae</i>
30	<i>Bifidobacterium primatium</i>	DSM 100687	Draft (8)	2696768	63.2	2073	2	57	51	2.5%	PEBI000000000	Cotton-top tamarin	<i>Saguinus oedipus</i>	<i>Cebidae</i>
31	<i>Bifidobacterium pseudocatenulatum</i>	DSM 20438	Complete	2313752	56.38	1816	6	54	65	3.6%	AP012330.1	Human	<i>Homo sapiens</i>	<i>Hominidae</i>
32	<i>Bifidobacterium ramosum</i>	DSM 100688	Draft (25)	3028921	63.51	2262	3	57	100	4.4%	WBSM000000000	Cotton-top tamarin	<i>Saguinus oedipus</i>	<i>Cebidae</i>
33	<i>Bifidobacterium reuteri</i>	DSM 23975	Draft (28)	2847572	60.45	2149	4	53	64	3.0%	JGZK000000000	Common marmoset	<i>Callithrix jacchus</i>	<i>Cebidae</i>
34	<i>Bifidobacterium saguini</i>	DSM 23967	Draft (33)	2787036	56.35	2321	5	59	72	3.1%	JGZN000000000	Red-handed tamarin	<i>Saguinus midas</i>	<i>Cebidae</i>
35	<i>Bifidobacterium samirii</i>	LMG 30940	Draft (42)	2574625	66.64	1968	5	60	91	4.6%	QXGK000000000	Black-capped squirrel monkey	<i>Saimiri boliviensis</i>	<i>Cebidae</i>
36	<i>Bifidobacterium scaligerum</i>	DSM 103140	Draft (19)	2650899	58.27	2063	2	59	59	2.9%	PGLQ000000000	Cotton-top tamarin	<i>Saguinus oedipus</i>	<i>Cebidae</i>
37	<i>Bifidobacterium scardovii</i>	DSM 13734	Complete	3158347	64.63	2489	3	56	117	4.7%	AP012331.1	Human	<i>Homo sapiens</i>	<i>Hominidae</i>
38	<i>Bifidobacterium simiarum</i>	DSM 103153	Draft (19)	2714697	63.79	2083	1	58	64	3.1%	PEBK000000000	Emperor tamarin	<i>Saguinus imperator</i>	<i>Cebidae</i>
39	<i>Bifidobacterium Stellenboschense</i>	DSM 23968	Draft (40)	2812864	65.34	2202	6	59	46	2.1%	JGZP000000000	Red-handed tamarin	<i>Saguinus midas</i>	<i>Cebidae</i>
40	<i>Bifidobacterium tissieri</i>	DSM 100201	Draft (38)	2873483	61.05	2255	2	60	58	2.6%	MWWV000000000	Common marmoset	<i>Callithrix jacchus</i>	<i>Cebidae</i>
41	<i>Bifidobacterium vansinderenii</i>	LMG 30126	Draft (68)	3111005	62.47	2515	4	60	66	2.6%	NEWD000000000	Emperor tamarin	<i>Saguinus imperator</i>	<i>Cebidae</i>

*Predicted number of rRNA loci based on draft genome sequences