Journal Pre-proof

Succession of *Bifidobacterium longum* strains in response to a changing early life nutritional environment reveals dietary substrate adaptations

Magdalena Kujawska, Sabina Leanti La Rosa, Laure C. Roger, Phillip B. Pope, Lesley Hoyles, Anne L. McCartney, Lindsay J. Hall

PII: S2589-0042(20)30556-3

DOI: https://doi.org/10.1016/j.isci.2020.101368

Reference: ISCI 101368

To appear in: *ISCIENCE*

Received Date: 11 March 2020

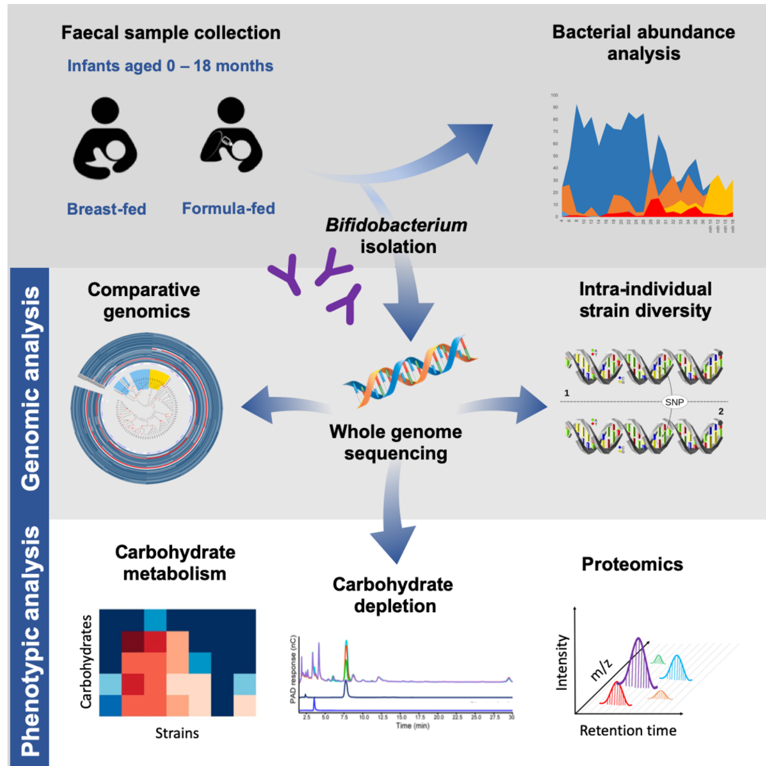
Revised Date: 16 June 2020

Accepted Date: 13 July 2020

Please cite this article as: Kujawska, M., La Rosa, S.L., Roger, L.C., Pope, P.B., Hoyles, L., McCartney, A.L., Hall, L.J., Succession of *Bifidobacterium longum* strains in response to a changing early life nutritional environment reveals dietary substrate adaptations, *ISCIENCE* (2020), doi: https://doi.org/10.1016/j.isci.2020.101368.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. 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.

© 2020



Journal Pre-proof

Journal Pre-proof

1 **Title**

2

3 **Succession of *Bifidobacterium longum* strains in response to a changing early life**
4 **nutritional environment reveals dietary substrate adaptations**

5

6 **Author list and affiliations**

7

8 **Magdalena Kujawska[1], Sabina Leanti La Rosa[2], Laure C. Roger[5], Phillip B. Pope[2,3], Lesley**
9 **Hoyles[4], Anne L. McCartney[5], Lindsay J. Hall[1,6,7,8,9,*]**
10 [1] Gut Microbes & Health, Quadram Institute Biosciences, Norwich Research Park, Norwich,
11 NR4 7UQ, UK
12 [2] Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life
13 Sciences, 1433, Aas, Norway
14 [3] Faculty of Biosciences, Norwegian University of Life Sciences, 1433, Aas, Norway
15 [4] Department of Biosciences, Nottingham Trent University, Nottingham, NG11 8NS, UK
16 [5] Department of Food & Nutritional Sciences, University of Reading, Reading, RG6 6LA, UK
17 [6] Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich, NR4
18 7TJ, UK
19 [7] Chair of Intestinal Microbiome, School of Life Sciences, Technical University of Munich,
20 85354 Freising, Germany
21 [8] ZIEL – Institute for Food & Health, Technical University of Munich, 85354 Freising, Germany
22 [9] Lead Contact

23

24 **Lead Contact info**

25

26 * Correspondence: Lindsay.Hall@quadram.ac.uk

27

28 **Summary**

29

30 Diet-microbe interactions play a crucial role in modulation of the early life microbiota and
31 infant health. *Bifidobacterium* dominates the breast-fed infant gut and may persist in

Journal Pre-proof

32 individuals during transition from a milk-based to a more diversified diet. Here, we
33 investigated adaptation of *B. longum* to the changing nutritional environment. Genomic
34 characterisation of 75 strains isolated from nine either exclusively breast- or formula-fed
35 (pre-weaning) infants in their first 18 months revealed subspecies- and strain-specific intra-
36 individual genomic diversity with respect to carbohydrate metabolism, which corresponded
37 to different dietary stages. Complementary phenotypic studies indicated strain-specific
38 differences in utilisation of human milk oligosaccharides and plant carbohydrates, while
39 proteomic profiling identified gene clusters involved in metabolism of selected
40 carbohydrates. Our results indicate a strong link between infant diet and *B. longum* diversity
41 and provide additional insights into possible competitive advantage mechanisms of this
42 *Bifidobacterium* species and its persistence in a single host.
43

44 **Key words**: *Bifidobacterium longum*, infant diet, carbohydrates, genomics, proteomics
45

46 **Introduction**

47 Microbial colonisation shortly after birth is the first step in establishment of the mutualistic
48 relationship between the host and its microbiota (Backhed et al., 2015, Wampach et al.,
49 2017, Lawson et al., 2020). The microbiota plays a central role in infant development by
50 modulating immune responses, providing resistance to pathogens, and also digesting the
51 early life diet (Heikkila and Saris, 2003, Sela et al., 2008, Marcobal and Sonnenburg, 2012,
52 Sivan et al., 2015, de Aguero et al., 2016, Thongaram et al., 2017). Indeed, diet-microbe
53 interactions are proposed to play a crucial role during infancy and exert health effects that
54 extend to later life stages (Turnbaugh et al., 2006, Renz et al., 2012, Olszak et al., 2012, Feng
55 et al., 2015, Bokulich et al., 2016, Tang et al., 2017). The gastrointestinal tract of vaginally
56 delivered full-term healthy infants harbours a relatively simple microbiota characterised by
57 the dominance of the genus *Bifidobacterium* (Dogra et al., 2015, Shao et al., 2019). In
58 contrast, caesarean section born infants have disrupted transmission of
59 maternal gastrointestinal bacteria, such as *Bifidobacterium*, and high levels of opportunistic
60 hospital-associated pathogens (Shao et al., 2019).
61 Breast milk is considered the gold nutritional standard for infants, which also acts as an
62 important dietary supplement for early life microbial communities, including
63 *Bifidobacterium*. The strong diet-microbe association has further been supported by reports

64 of differences in microbial composition between breast- and formula-fed infants (e.g. high
65 versus low *Bifidobacterium* abundance) and related differential health outcomes between
66 the two groups: e.g. increased instances of asthma, allergy and obesity in formula-fed
67 infants (Ip et al., 2007, Das, 2007, O'Sullivan et al., 2015, Martin et al., 2016, Stiemsma and
68 Michels, 2018, Ortega-Garcia et al., 2018, Forbes et al., 2018).
69 The high abundance of *Bifidobacterium* in breast-fed infants has been linked to the presence
70 of specific carbohydrate utilisation genes and gene clusters in their genomes, particularly
71 the ones involved in the degradation of breast milk-associated human milk oligosaccharides
72 (HMOs) (Sela et al., 2008). The presence of these genes is often species- and indeed strain-
73 specific, and has been described in *B. breve*, *B. bifidum*, *B. longum*, *B. infantis*, and more
74 rarely in *B. pseudocatenulatum* (Sela et al., 2008, James et al., 2016, Katayama, 2016,
75 Garrido et al., 2016). However, previous studies have indicated co-existence of
76 *Bifidobacterium* species and strains in individual hosts, resulting in interaction and metabolic
77 co-operation within a single (HMO-associated) ecosystem (Milani et al., 2015a, Lawson et
78 al., 2020).
79 Transition from breastfeeding to a more diversified diet and the introduction of solid foods
80 has been considered to initiate the development of a functionally more complex adult-like
81 microbiome with genes responsible for degradation of plant-derived complex
82 carbohydrates, starches, and xenobiotics, as well as production of vitamins (Koenig et al.,
83 2011, McKeen et al., 2019). Non-digestible complex carbohydrates such as inulin-type
84 fructans (ITF), arabino-xylans (AX) or arabinoxylo-oligosaccharides (AXOS) in complementary
85 foods have been proposed to potentially exert beneficial health effects through their
86 bifidogenic and prebiotic properties and resulting modulation of the intestinal microbiota
87 and metabolic end-products (Roberfroid, 2007, Broekaert et al., 2011, Hald et al., 2016,
88 Riviere et al., 2016).
89 Despite the shift in microbiota composition during weaning, specific strains of
90 *Bifidobacterium*, and *B. longum* in particular, have previously been shown to persist in
91 individuals over time (Maldonado-Gomez et al., 2016, Oki et al., 2018). *B. longum* is
92 currently recognised as four subspecies: *longum* and *infantis* (characteristic of the human
93 gut microbiota), and *suis* and *suillum* (from animal hosts) (Mattarelli et al., 2008, Yanokura
94 et al., 2015). It is considered the most common and prevalent species found in the human
95 gut, with *B. longum* subsp. *infantis* detected in infants, and *B. longum* subsp. *longum* widely

Journal Pre-proof

Journal Pre-proof
96 distributed in both infants and adults (Turroni et al., 2009, Turroni et al., 2012). The
97 differences in prevalence between the two subspecies, and the ability of infant, adult and
98 elderly host to acquire new *B. longum* strains during a lifetime have been attributed to
99 distinct bacterial carbohydrate utilisation capabilities and the overall composition of the
100 resident microbiota (Garrido et al., 2012, Odamaki et al., 2018).
101 There have been several recent studies that have explored the early life microbiota in
102 breast- and formula-fed babies (Magne et al., 2006, Palmer et al., 2007, Roger and
103 McCartney, 2010, Roger et al., 2010). Strain-level metagenomic investigation of the
104 DIABIMMUNE cohort provided insights into diet-related functional aspects of *B. infantis* in
105 breast-fed infants (Vatanen et al., 2019). Longitudinal studies focusing specifically on *B.*
106 *longum* have highlighted intraspecies diversity, colonisation and long-term persistence
107 (years) of this species in hosts, however there have been limited investigations into diet-
108 related functions at early life stages (Chaplin et al., 2015, Oki et al., 2018, Odamaki et al.,
109 2018). Furthermore, although there are studies examining *B. longum* strains in relation to
110 diet, these have not been profiled over longitudinal and changing dietary periods (Arboleya
111 et al., 2018). Hence, longitudinal assessments of *B. longum* strains in single hosts over time,
112 and with focus on changing dietary patterns are lacking, and further detailed studies are
113 required.
114 Here, we investigate the adaptations of *Bifidobacterium* to the changing infant diet and
115 examine a unique collection of *B. longum* strains isolated from nine infants across their first
116 18 months, encompassing pre-weaning, weaning and post-weaning phases. We probed the
117 genomic and phenotypic similarities between 62 *B. longum* strains and 13 *B. infantis* strains
118 isolated from either exclusively breast-fed or formula-fed infants (pre-weaning). Our results
119 indicate a strong link between host diet and *Bifidobacterium* species/strains, which appears
120 to correspond to the changing nutritional environment.
121
122 **Results**
123 Previous investigations into *B. longum* across the human lifespan have determined a broad
124 distribution of this species, including prolonged periods of colonisation (Maldonado-Gomez
125 et al., 2016, Oki et al., 2018). To gain insight into potential mechanisms facilitating these
126 properties during the early life window, we investigated the genotypic and phenotypic
127 characteristics of *B. longum* strains within individual infant hosts in relation to diet (i.e.

Journal Pre-proof

128 breast milk vs. formula) and dietary stages (i.e. pre-weaning, weaning and post-weaning),
129 following up on a longitudinal study of the infant faecal microbiota published in 2010 (Roger
130 and McCartney, 2010). Briefly, faecal samples from exclusively breast-fed infants and
131 exclusively formula-fed infants were collected regularly from 1 month to 18 months of age
132 (Roger and McCartney, 2010). The number of samples obtained from the breast-fed infants
133 during the pre-weaning period was higher than that obtained from the formula-fed group,
134 which may correlate with differences in weaning age (~20.6 vs. ~17 weeks old). Collected
135 samples were subjected to quantitative analysis using fluorescence in situ hybridization
136 (FISH) to enumerate the predominant bacterial groups (**Table S1**) (Roger and McCartney,
137 2010). Bacterial isolation was also carried out on selected samples, and the isolated colonies
138 identified using ribosomal intergenic spacer analysis (Roger and McCartney, 2010).
139
140 **Quantitative analysis of microbial communities in breast- and formula-fed infants**
141 To provide context to the microbiome environment the strains selected for the present
142 study were isolated from, we re-analysed the data originally generated by FISH (**Figure 1,**
143 **Table S1**) (Roger and McCartney, 2010). Bacteria detected using probe Bif164
144 (bifidobacteria) proportionally constituted the predominant group in samples isolated from
145 breast-fed infants during pre-weaning and weaning; between 16.5% to 100% of the
146 microbiota across the study period. During post-weaning, proportions of bifidobacteria
147 across all breast-fed samples decreased considerably and ranged from 4.6% to 12.1%. The
148 levels of bacteria detected by ER482 (members of *Clostridium* cluster XIVa) started to
149 increase during weaning, increasing to 18.2% (from 0.25% at pre-weaning). Bacteria
150 detectable by probe Bac303 (members of genus *Bacteroides*, *Parabacteroides* and *Prevotella*
151 species, *Paraprevotella*, *Xylanibacter*, *Barnesiella* species and *Odoribacter splanchnicus*)
152 were identified in all samples throughout the study, with this bacterial group showing
153 extensive inter-individual variation. Other microbiota members were detected in breast-fed
154 samples at lower levels, including members of family Coriobacteriia (Ato291, mean <2% of
155 microbiota), *Escherichia coli* (EC1531, <1%), members of *Clostridium* clusters I and II
156 (Chis150, <1%) and lactic acid bacteria (Lab158, mean <1%).
157 In contrast to the breast-fed group, no drastic shift in bacterial populations was observed in
158 formula-fed infants throughout the study. Overall, lower levels of bifidobacteria were
159 detected during pre-weaning and weaning, fluctuating from 0.0% to 73.3% of the

Journal Pre-proof

160 microbiota at different time points. Similar to the breast-fed group, proportions of
161 *Bifidobacterium* decreased during post-weaning across all formula-fed samples and ranged
162 from 6.5% to 12% at month 18. The levels of bacteria detected by probe ER482 were overall
163 higher in formula-fed samples throughout study duration; 19.96±17.41%, 25.39±14.63%
164 and 30.6±15.92% for pre-weaning, weaning and post-weaning phases. Similarly, proportions
165 of bacteria detected by Bac303 during all dietary phases were higher in the formula-fed
166 group compared to the breast-fed group. Contrastingly to the breast-fed group, levels of
167 bacteria detected by Chis150 (*Clostridium* clusters I and II) started to increase during
168 weaning in the formula-fed group and continued to increase (1.23±1.28%, 7.03±9.18% and
169 21.72±11.47% for pre-weaning, weaning and post-weaning, respectively). Levels of bacteria
170 identified by Ato291 and EC1531 in formula-fed samples were slightly higher than in the
171 breast-fed group (means of <3.5% and <1.25%, respectively), while the mean proportion of
172 lactic acid bacteria (Lab158) remained below <1%.
173 Overall, these results confirm previous studies which have indicated differences in faecal
174 microbiota composition between breast- and formula-fed babies, particularly during the
175 pre-weaning and weaning phases, and demonstrate the succession of bacterial species over
176 time and in relation to diet, including *Bifidobacterium*.
177

178 **General features of *B. longum* genomes**

179 Based on the results of bacterial culture and colony identification published previously (for
180 details, refer to (Roger et al., 2010)), 88 isolates originally identified as *Bifidobacterium* were
181 selected for this study, 46 from five exclusively breast-fed infants (BF1-BF5, including
182 identical twins BF3 and BF4) and 42 from four exclusively formula-fed infants (FF1-FF3 and
183 FF5). Following sequencing and ANI analysis (**Tables S2 & S3**), 75 strains were identified as
184 *B. longum* sp. and included in further analysis, with 62 strains identified as *B. longum* subsp.
185 *longum* (*B. longum*) and 13 strains identified as *B. longum* subsp. *infantis* (*B. infantis*) (**Figure**
186 **2a**).
187 To determine possible genotypic factors facilitating establishment and persistence of *B.*
188 *longum* in the changing early life environment, we assessed the genome diversity of our
189 strains. Sequencing generated between 12 and 193 contigs for each *B. longum* strain, with
190 74/75 containing fewer than 70 contigs, yielding a mean of 66.95-fold coverage for strains

191 (**Table S2**). The predicted genome size for strains identified as *B. longum* ranged from 2.21
192 Mb to 2.58 Mb, possessing an average G+C% content of 60.11%, an average predicted ORF
193 number of 2,023 and number of tRNA genes ranging from 55-88. For strains identified as *B.*
194 *infantis*, the predicted genome size ranged from 2.51 Mb to 2.75 Mb, with an average G+C%
195 content of 59.69%, an average predicted ORF number of 2,280 and the number of tRNA
196 genes ranging from 57 to 62.
197
198 **Comparative genomics**
199 To identify *B. longum* strains among the sequenced isolates and assess nucleotide-level
200 genomic differences, we performed ANI analysis. Results (**Table S3**) indicated that *B. longum*
201 strains isolated from individual infant hosts displayed higher levels of sequence identity than
202 strains isolated from different hosts. More specifically, pairwise identity values for strains
203 isolated from infant BF3 showed the narrowest range (average value of 99.99±3.15e-5%),
204 followed by infant FF2 strains (99.98±1.12e-4%), with infant BF2 strains having the broadest
205 identity value range (averaging 99.13±7.8e-3%).
206 Next, we examined genetic diversity of newly sequenced *B. longum* strains and their
207 relatedness to each other, alongside *B. longum* type strains. We identified a total of 1002
208 core genes present in at least 99% of the analysed *B. longum* subspecies genomes that
209 allowed clear distinction between *B. longum* subspecies (i.e. *longum* vs. *infantis*) based on
210 the presence/absence of specific genes (**Table S4**). Phylogenetic analysis revealed that *B.*
211 *longum* strains within each subspecies clustered mainly according to isolation source, i.e.
212 individual infants, rather than dietary stage (i.e. pre-weaning, weaning and post-weaning)
213 (**Figure 2b**). Interestingly, strains isolated from formula-fed baby FF5 clustered into two
214 separate clusters, irrespective of the isolation period, suggesting presence of two highly
215 related *B. longum* groups within this infant. Furthermore, strains isolated from identical
216 twins BF3 and BF4 clustered together, indicating their close relatedness.
217 We next sought to identify whether specific components of the *B. longum* subspecies
218 pangenome were enriched in infant hosts. Each candidate gene in the accessory genome
219 was sequentially scored according to its apparent correlation to host diet (breast vs.
220 formula) or dietary stage. A gene annotated as α-L-arabinofuranosidase, along with four
221 other genes coding for hypothetical proteins, were predicted to be enriched in *B. longum*
222 strains isolated from breast-fed infants. Alpha-L-arabinofuranosidases are enzymes involved

Journal Pre-proof

223 in hydrolysis of terminal non-reducing α-L-arabinofuranoside residues in α-L-arabinosides
224 and act on such carbohydrates as (arabino)xylans (Ichinose et al., 2008, Ahmed et al., 2013).
225 In addition, two genes coding for hypothetical proteins and a gene coding for Mobility
226 protein A were overrepresented in strains isolated from formula-fed infants. Furthermore,
227 no associations between genes and dietary stages in *B. longum*, nor any associations
228 whatsoever in *B. infantis* were observed (**Table S5**).
229 As our strains were isolated from individual infants at different time points, we next sought
230 to determine their intra-strain diversity; for this we used the first *B. longum* isolate from
231 each infant as the 'reference' strain to which all other strains from the same infant were
232 compared (**Figure 3**). Infants BF1, BF3 and FF2 had the lowest strain diversity; with
233 respective mean pairwise SNP distances of 18.7±20.3 SNPs (mean±sd), 10.3±5.0 SNPs and
234 13.3±5.3 SNPs. These results suggest strains isolated from these infants may be clonal,
235 indicating long-term persistence despite dietary changes. Surprisingly, analysis of strains
236 isolated from breast-fed identical twins BF3 and BF4 revealed higher strain diversity in baby
237 BF4 (1034.5±1327.1 SNPs), compared to the highly similar strains in infant BF3 (i.e. 10.3±5.0
238 SNPs). Based on these results, we conducted SNP analysis on *B. longum* strains isolated from
239 both babies and found that out of 13 strains analysed (n=8 from BF3 and n=5 from BF4), 12
240 isolated during pre-weaning, weaning and post-weaning appeared to be clonal (with mean
241 pairwise SNP distance of 10.0±5.5 SNPs) and one strain from baby BF4 isolated post-
242 weaning was more distant, 2595.4±2.8 SNPs. The difference in strain diversity may relate to
243 the fact that infant BF4 received a course of antibiotics during pre-weaning (**Figure1, Tables**
244 **S1 & S2**) (Roger and McCartney, 2010). Furthermore, the presence of clonal strains in both
245 babies suggests vertical transmission of *B. longum* from mother to both infants, or potential
246 horizontal transmission between babies, consistent with previous reports (Makino et al.,
247 2011, Makino et al., 2013, Milani et al., 2015b, Odamaki et al., 2018). *B. infantis* strains
248 isolated from infant BF2 showed the highest strain diversity, 9030.9±8036.6 SNPs. Seven
249 strains isolated during both pre-weaning and weaning periods appeared to be clonal,
250 6.3±1.6 SNPs, while four strains isolated during weaning and post-weaning were more
251 distant, with mean pairwise SNP distance of 14983.5±4658.3 SNPs (**Table S6**).
252
253 **Functional annotation of *B. longum* subspecies genomes – carbohydrate utilisation**

254 To assess genomic differences between our strains at a functional level, we next assigned
255 functional categories to ORFs of each *B. longum* genome. Carbohydrate transport and
256 metabolism was identified as the second most abundant category (after unknown function),
257 reflecting the saccharolytic lifestyle of *Bifidobacterium* (**Figure S1**) (Pokusaeva et al., 2011,
258 Milani et al., 2015a). *B. longum* had a slightly higher proportion of carbohydrate metabolism
259 and transport genes (11.39±0.31%) compared to *B. infantis* (10.20±0.60%), which is
260 consistent with previous reports (Ventura et al., 2009, Sela and Mills, 2010). *B. longum*
261 strains isolated during pre-weaning had a similar proportion of carbohydrate metabolism
262 genes in comparison with the strains isolated post-weaning: 11.28±0.23% and 11.48±0.38%,
263 respectively. Furthermore, we obtained similar results for *B. longum* strains isolated from
264 breast- and formula-fed infants, with respective values of 11.41±0.21% and 11.38±0.38%. In
265 contrast, *B. infantis* strains isolated pre-weaning had a lower proportion of carbohydrate
266 metabolism genes in their genomes compared to the ones isolated post-weaning:
267 9.90±0.24% and 11.20±0.01%, respectively (**Table S7**).
268 One of the major classes of carbohydrate-active enzymes comprises glycosyl hydrolases
269 (GH), which facilitate glycan metabolism in the gastrointestinal tract. We thus sought to
270 investigate and compare the arsenal of GHs in *B. longum* using dbCAN2. We identified a
271 total of 36 different GH families in all *Bifidobacterium* strains. *B. longum* was predicted to
272 contain 55 GH genes per genome on average (2.72 % of OFRs), while this number was lower
273 for *B. infantis* strains, ~37 GH genes per genome (1.62% of ORFs) (**Figure 4, Table S8**). The
274 predominant GH family was GH43; enzymes involved in metabolism of complex plant
275 carbohydrates such as (arabino)xylans (Viborg et al., 2013), followed by GH13 (starch), GH51
276 (hemicelluloses) and GH3 (plant glycans) (Milani et al., 2015a, Milani et al., 2016).
277 Within the *B. longum* group, strains isolated during pre-weaning had a slightly lower mean
278 number of GH genes compared to strains isolated post-weaning (54.46±2.81 vs.
279 56.85±2.77). Moreover, strains isolated from breast-fed babies contained an average of
280 53.96±3.82 GH genes per genome, while this number was slightly higher for strains isolated
281 from formula-fed infants; 56.47±2.96. Further analysis revealed that these differences
282 appeared to be intra-host-specific and diet-related. For example, strains isolated from
283 breast-fed twins BF3 and BF4 pre-weaning had 11 GH43 genes per genome, while the pre-
284 weaning strain from formula-fed baby FF3 had 13 GH genes per genome predicted to belong
285 to this GH family. Similarly, strains isolated from babies BF3 and BF4 post-weaning had 11

Journal Pre-proof

Journal Pre-proof

286 predicted GH genes, while the three strains isolated from infant FF3 were predicted to
287 contain 16, 16 and 18 GH genes per genome, respectively (**Table S8**).
288 We next determined if these GH genes differences statistically correlated with breast- and
289 formula-fed groups (**Table S8**). Significant differences ($p < 0.05$) were observed between
290 mean numbers of GH genes belonging to the predominant GH families (GH43 – higher
291 abundance in FF babies, GH13 – higher abundance in BF babies, and GH51 – higher
292 abundance in FF babies), and several other GH families, including GH5 (β-glucosidases and
293 β-mannosidases), GH38 (mannosylglycerate hydrolases) and GH36 (α-galactosidases), all
294 more abundant in BF babies. Further analysis of dietary phases suggested significant
295 differences in GH genes between breast- and formula-fed groups during pre-weaning (e.g.
296 families GH43, GH13, GH5, GH38), but not in the post-weaning phase (**Table S8**).
297 Since glycosyl hydrolases belonging to distinct GH families may have similar catalytic
298 properties, we next grouped the GH genes for which the predicted enzyme class annotation
299 was available and investigated their abundance (**Table S9**). The predominant enzyme classes
300 in *B. longum* strains were non-reducing end α-L-arabinofuranosidases belonging to GH43
301 and GH51, followed by β-galactosidases (GH2 and GH42), oligo-1,6-glucosidases (GH13) and
302 β-*N*-acetylhexosaminidases (GH3 and GH20).
303 The mean numbers of enzyme classes between breast- and formula-fed babies significantly
304 differed ($p < 0.05$) in the top three above-mentioned predominant enzyme classes as well as
305 several other less abundant ones, including non-reducing end β-L-arabinofuranosidases
306 (GH127 and GH146 – higher abundance in BF babies), α-galactosidases (GH36 – higher
307 abundance in BF babies), and endo-1,5-α-L-arabinases (GH43 – higher abundance in FF
308 babies). Additional analysis of dietary phases indicated significant differences between
309 breast- and formula fed groups during pre-weaning (e.g. non-reducing end α-L-
310 arabinofuranosidases, β-galactosidases, oligo-1,6-glucosidases as well as α-galactosidases),
311 but not during post-weaning (**Table S9**).
312 We next examined the predicted glycosyl hydrolase repertoire of *B. infantis* strains, with the
313 caveat that the majority of the strains belonging to this subspecies were isolated from a
314 single infant. In contrast to the *B. longum* group, the most abundant GH family was GH13
315 (starch), followed by GH42, GH20 and GH38 (**Table S8**). *B. infantis* strains also harboured
316 genes predicted to encode members of the GH33 family, which contains exo-sialidases

Journal Pre-proof

317 (Milani et al., 2015a). Strains isolated pre-weaning were predicted to contain an average of
318 34.83±0.4 GH genes per genome, while this number was higher for the strains isolated post-
319 weaning (i.e. 43.00±0.00 GH genes). *B. infantis* strains isolated post-weaning contained
320 families GH1 and GH43 that were absent in the strains isolated pre-weaning. The GH1 family
321 contains enzymes such as β-glucosidases, β-galactosidases and β-D-fucosidases active on a
322 wide variety of (phosphorylated) disaccharides, oligosaccharides, and sugar–aromatic
323 conjugates (Suzuki et al., 2013). The analysis of enzyme classes in the *B. infantis* strains
324 suggested that β-galactosidases (GH2 and GH42) were predominant in this group, followed
325 by β-*N*-acetylhexaminidases (GH3 and GH20), 4-α-glucanotransferases (GH77) and oligo-
326 1,6-glucosidases (GH13) (**Table S9**).
327 Members of the genus *Bifidobacterium* have previously been shown to contain GH genes
328 involved in metabolism of various HMOs present in breast milk (Garrido et al., 2015, Garrido
329 et al., 2016). We identified genes belonging to GH29 and GH95 (α-L-fucosidases found
330 active on fucosylated HMOs (Sela et al., 2012, Garrido et al., 2016)) in all our *B. infantis*
331 strains, as well as four *B. longum* strains isolated from formula-fed baby FF3. Furthermore,
332 we found GH20 and GH112 genes (lacto-*N*-biosidases and galacto-*N*-biose/lacto-*N*-biose
333 phosphorylases shown to be involved in degradation of isomeric lacto-*N*-tetraose (LNT)
334 (Kitaoka, 2012)) in all our *B. infantis* and *B. longum* strains (**Table S8**).
335 Overall, these findings suggest differences in general carbohydrate utilisation at different
336 stages suggesting adaptation of *Bifidobacterium* to a changing early life nutritional diet,
337 which may be a factor facilitating establishment of these bacteria within individuals during
338 infancy.
339

340 **Prediction of gain and loss of GH families in *B. longum***

341 Given the differences in the carbohydrate utilisation profiles between *B. longum* and *B.*
342 *infantis*, we next investigated the acquisition and loss of GH families. For this purpose, we
343 additionally predicted the presence of GH families in type strains *B. longum* subsp. *longum*
344 JCM 1217^T, *B. longum* subsp. *infantis* ATCC 15697^T and *B. longum* subsp. *suis* LMG 21814^T
345 with dbCAN2 and generated a whole genome SNP tree to reflect gene loss/gain events more
346 accurately (**Figure 4, Table S10**). Both *B. longum* and *B. infantis* lineages appear to have
347 acquired GH families (when compared to the common ancestor of the phylogenetic group),

348 with the *B. longum* lineage gaining two GH families (GH121 and GH146) and the *B. infantis*
349 lineage one GH family (GH33). Within the *B. infantis* lineage, which also contains the *B. suis*
350 type strain, the *B. infantis* taxon has further acquired two and lost five GH families. These
351 findings suggest that the two human-related subspecies have followed different
352 evolutionary paths, which is consistent with our observation of differences between *B.*
353 *longum* and *B. infantis* resulting from phylogenomic analyses. Intriguingly, strain adaptation
354 to the changing host environment (i.e. individual infant gut) seems to be driven by loss of
355 specific GH families (**Figure 4**). For example, *B. infantis* strains isolated during pre-weaning
356 and weaning from baby BF2 appear to be missing up to three GH families (GH1, GH43 and
357 GH109) present in strains isolated post-weaning. Lack of family GH43 (containing enzymes
358 involved in metabolism of a variety of complex carbohydrates, including plant-derived
359 polysaccharides) in early life *B. infantis* strains may explain nutritional preference of this
360 subspecies for an HMO-rich diet. Similarly, we observed differential gene loss events in *B.*
361 *longum* strains from individual hosts. For example, all strains isolated from baby BF5 appear
362 to lack GH families GH1, GH29 and GH95. However, strains isolated pre-weaning
363 additionally lacked GH53 family, which includes endogalactanases shown to be involved in
364 liberating galactotriose from type I arabinogalactans in *B. longum* (Hinz et al., 2005). In
365 contrast, strain B_38 isolated from this infant (BF5) post-weaning appears to have lost
366 families GH136 and GH146. Interestingly, members of family GH136 are lacto-*N*-biosidases
367 responsible for liberating lacto-*N*-biose I from LNT, an abundant HMO unique to human milk
368 (Yamada et al., 2017), while family GH146 contains β-L-arabinofuranosidases displaying <u>*exo*</u>-
369 activity on β-linked arabinofuranosyl groups. These events may be linked to dietary changes
370 (withdrawal of breast milk) and/or a shift in the composition of the microbiota post-
371 weaning. Only one *B. longum* strain was isolated post-weaning from this baby, however
372 FISH analysis (**Figure 1, Table S1**) revealed an increase in the bacteroides group, which might
373 explain the loss of family GH146 by strain B_38 as the founding member of GH146 family, β-
374 L-arabinofuranosidase, was first characterised in *Bacteroides thetaiotaomicron* (Luis et al.,
375 2018). Overall, the presence of intra-individual and strain-specific GH family repertoires in *B.*
376 *longum* suggests their adaptation to host-specific diet. The presence of strains with different
377 GH content at different dietary stages further indicates potential acquisition of new
378 *Bifidobacterium* strains with nutrient-specific adaptations in response to the changing infant
379 diet.

Journal Pre-proof

380

381 **Prediction of single nucleotide polymorphisms (SNPs) in glycosyl hydrolases**

382 Given the intra-strain diversity in the nine babies and the differences in GH repertoires
383 between *B. longum* and *B. infantis*, we next sought to examine nucleotide-level differences
384 in glycosyl hydrolase genes between strains in individual infants (**Table S11**). Unsurprisingly,
385 we did not identify any significant SNPs that may lead to functional changes in GH genes in
386 infants that had the lowest strain diversity (infants BF1, BF3 and FF2) (**Table S6**). The highest
387 number of GH genes with predicted variants was recorded for *B. infantis* strains from baby
388 BF2. In total, 52 synonymous variants and 29 missense variants were predicted at 81
389 different positions in 12 GH genes across strains that showed the highest diversity from the
390 first 'reference' isolate, namely one strain isolated during weaning and the three strains
391 isolated post-weaning. A number of missense variants, both complex and single, were
392 recorded at several positions in the predominant enzyme classes, i.e. β-galactosidases (EC
393 3.2.1.23) and β-*N*-hexosaminidases (EC 3.2.1.52).
394 Similarly, both synonymous and missense variants were predicted in *B. longum* strains less
395 closely related to 'reference' strains from breast-fed (BF4 and BF5) and formula-fed (FF1,
396 FF3 and FF5) babies. We did not observe any trend in the distribution of SNPs across GH
397 genes in *B. longum* strains. The number of predicted variants, the number of GH genes with
398 identified mutations and their enzyme classification differed between individual infants. For
399 example, in baby BF4 9 out of 10 predicted variants (4 synonymous and 5 missense) were
400 identified in an α-xylosidase in a strain isolated post-weaning, while in baby FF5 14
401 synonymous and 10 missense variants were predicted at 24 positions in 7 different GH
402 genes across strains isolated during weaning and post-weaning. Some missense changes do
403 not compromise normal protein function, while others can change essential aspects of
404 protein maturation, activity or stability (Miosge et al., 2015). The presence of missense
405 variants in GH genes of *B. longum* strains may indicate potential functional differences and
406 provide additional explanation to intra-strain and intra-individual carbohydrate metabolism
407 profiles of these bacteria, however experimental evidence would be essential to confirm the
408 importance of these predictions.

409

410 **Phenotypic characterisation of carbohydrate utilisation**

Journal Pre-proof

411 *Bifidobacterium longum* has previously been shown to metabolise a range of carbohydrates,
412 including dietary and host-derived glycans (Watson et al., 2013, Arboleya et al., 2018). Given
413 the predicted differences in carbohydrate metabolism profiles, and to understand strain-
414 specific nutrient preferences, we next sought to determine their glycan fermentation
415 capabilities. We performed growth assays on 49 representative strains from all nine infants,
416 cultured in modified MRS supplemented with selected carbohydrates as the sole carbon
417 source. For these experiments, we chose both plant- and host-derived glycans that we
418 would expect to constitute components of the early life infant diet (Mills et al., 2019).
419 Although all *B. longum* strains were able to grow on simple carbohydrates (i.e. glucose and
420 lactose), we also observed subspecies-specific complex carbohydrate preferences,
421 consistent with bioinformatic predictions (**Figure 5**). To represent host-derived
422 carbohydrates, we selected 2'-fucosyllactose (2'-FL) and lacto-*N*-neotetraose (LNnT) as
423 examples of HMOs found in breast milk. Out of the tested isolates, all *B. infantis* strains
424 were able to metabolise 2'-FL, as were three *B. longum* strains isolated from a formula-fed
425 baby FF3 during weaning and post-weaning (**Figure 5**). These results supported the
426 computational analysis and the identification of genes potentially involved in degradation of
427 fucosylated carbohydrates in the genomes of these isolates (GH29 and GH95). Although
428 bioinformatics identified the presence of genes involved in metabolism of isomeric LNT in all
429 our strains (GH20 and GH112), LNnT metabolism in *B. infantis* was strain-specific, with most
430 strains showing what we considered moderate (above 0.15 difference in OD from time T_2) to
431 high growth rates (above 0.25 difference in OD from time T_2), with two strains displaying
432 inconsistent growth (**Table S12**). Out of *B. longum* strains, B_24 and B_25 (isolated during
433 weaning from breast-fed baby BF3) also showed growth on LNnT, albeit this was
434 inconsistent. In contrast to all other *B. longum* strains, strain B_25 was not able to
435 metabolise plant-derived arabinose and xylose despite the predicted presence of genes
436 involved in metabolism of monosaccharides (GH43, GH31, GH2). However, it was one of the
437 two strains (out of 49 tested) that showed growth on cellobiose in 2/3 experiments; the
438 other one being the post-weaning *B. infantis* strain B_19 isolated from baby BF2. Given
439 these interesting results, we performed additional assays using cellobiose as the sole carbon
440 source over 72h, in which the *B. longum* strain B_25 showed high growth rate (above 0.25
441 difference in OD from time T_2), while the *B. infantis* B_19 strain did not grow at all (**Table**
442 **S12**). Additionally, both *B. longum* and *B. infantis* strains showed varying degrees of growth

443 performance on mannose, even when analysing the same strain, while none of the tested
444 strains were able to grow on arabinogalactan, pectin or rhamnose (**Figure 5**).
445 To further characterise strains identified above for putative carbohydrate degradation
446 genes, we performed carbohydrate uptake analysis and proteomics. *B. longum* strain B_25,
447 from one of the breast-fed identical twins that showed growth on LNnT and cellobiose, and
448 formula-fed strain B_71 which was able to grow on 2'-FL, were chosen. Supernatant from
449 these cultures was initially subjected to high-performance anion-exchange chromatography
450 (HPAEC) to evaluate the carbohydrate-depletion profiles (**Figure 6**). In all three cases, the
451 chromatograms showed complete utilisation of the tested carbohydrates and absence of
452 any respective degradation products in the stationary phase culture. The depletion of
453 cellobiose by B_25 and 2'-FL by B_71 occurred in the early exponential phase while LNnT
454 was still detected in the culture supernatant until the late exponential phase of growth,
455 suggesting that cellobiose and 2'-FL were internalised more efficiently than LNnT. We next
456 determined the proteome of B_25 and B_71 when growing on cellobiose, LNnT and 2'-FL
457 compared to glucose (**Figure 6a-c & Table S13**). The top 10 most abundant proteins in the
458 cellobiose proteome of B_25 included three β-glucosidases belonging to GH3 family, as well
459 as a homologue of transport gene cluster previously shown to be upregulated in *B. animalis*
460 subsp. *lactis* Bl-04 during growth on cellobiose (**Figure 6a & Table S14**) (Andersen et al.,
461 2013). Among the three β-glucosidases, B_25_00240 showed 98% sequence identity to the
462 structurally characterized BlBG3 from *B. longum*, which has been shown to be involved in
463 metabolism of the natural glycosides saponins (Yan et al., 2018). B_25_01763 and
464 B_25_00262 showed 46% identity to the β-glucosidase Bgl3B from *Thermotoga neapolitana*
465 (Pozzo et al., 2010) and 83% identity to BaBgl3 from *B. adolescentis* ATCC 15703 (Florindo et
466 al., 2018), respectively, two enzymes previously shown to hydrolyse cello-oligosaccharides.
467 With respect to LNnT metabolism by the same strain, the most abundant proteins were
468 encoded by genes located in two gene clusters (B_25_00111-00117 and B_25_00130-
469 00133) with functions compatible with LNnT import, degradation to monosaccharides and
470 further metabolism. The gene clusters contain the components of an ABC-transporter
471 (B_25_00111-00113), a predicted intracellular GH112 lacto-*N*-biose phosphorylase
472 (B_25_00114), an *N*-acetylhexosamine 1-kinase (B_25_00115) and enzymes involved in the
473 Leloir pathway. All these proteins were close homologues to proteins previously implicated
474 in the degradation of LNT/LNnT by type strain *B. infantis* ATCC 15697^{T} (Ozcan and Sela,

Journal Pre-proof

475 2018) (**Figure 6b & Table S14**). Interestingly, all clonal strains isolated from twin babies BF3
476 and BF4 also contained close homologues of all the above-mentioned genes in their
477 genomes, in some cases identical to those determined in B_25; however, only strain B_25
478 was able to grow on cellobiose and LNnT. Growth of B_71 on 2'-FL corresponded to
479 increased abundance of proteins encoded by the gene cluster B_71_00973-00983. These
480 proteins showed close homology to proteins described for *B. longum* SC596 and included
481 genes for import of fucosylated oligosaccharides, fucose metabolism and two α-fucosidases
482 belonging to the families GH29 and GH95 (**Figure 6c & Table S14**) (Garrido et al., 2016).
483
484 **Discussion**
485 High abundance of *Bifidobacterium,* and *B. longum* in particular, in early infancy is strongly
486 linked to availability of nutrients (Koenig et al., 2011, Backhed et al., 2015, Yamada et al.,
487 2017). In this study, we aimed to investigate the adaptations of *B. longum* to the changing
488 infant diet during the early life developmental window. Profiling microbiota composition
489 data (Roger and McCartney, 2010), genomic diversity of 75 *B. longum* strains (isolated from
490 infants at different dietary stages), and growth performance on different carbohydrates
491 revealed intra-individual and diet-related differences, which links to strain-level metabolism
492 properties for specific dietary components.
493 The FISH results corroborate findings of previous studies investigating the infant gut
494 microbiota; inter-individual variability during pre-weaning and weaning, with a shift towards
495 a more adult-like faecal microbiota associated with more complex diet at post-weaning
496 across all samples (Koenig et al., 2011, McKeen et al., 2019). *Bifidobacterium* constituted the
497 predominant group in breast-fed infants during pre-weaning and weaning, while the
498 composition of microbiota of the formula-fed infants during these stages was more
499 complex.
500 Our comparative genomic analysis indicates that clonal strains of *B. longum* can persist in
501 individuals through infancy, for at least 18 months, despite significant changes in diet during
502 weaning, which is consistent with previous reports (Maldonado-Gomez et al., 2016,
503 Odamaki et al., 2018). Concurrently, new strains (that display different genomic content and
504 potential carbohydrate metabolism capabilities) can be acquired, possibly in response to the
505 changing diet. Previously, strain shift in relation to withdrawal of breast milk has been
506 suggested for *B. infantis* by Vatanen et al. (Vatanen et al., 2019) based on strain-level

Journal Pre-proof

507 metagenomic approach. Similarly, Asnicar et al. (Asnicar et al., 2017) suggested that
508 originally acquired maternal strains of *B. longum* can be replaced at later life stages. Initial
509 vertical acquisition of *Bifidobacterium* from mother to newborn babies has been well
510 documented (Mikami et al., 2012, Makino et al., 2013, Milani et al., 2015b, Asnicar et al.,
511 2017); however, details of strain transmission events in later life are currently unclear. Work
512 of Odamaki et al. (Odamaki et al., 2018) suggested person-to-person horizontal transmission
513 of a particular *B. longum* strain between members of the same family, with direct transfer,
514 common dietary sources or environmental reservoirs, such as family homes (Lax et al.,
515 2014), as potential vehicles and routes for strain transmission. Our results showed the
516 presence of clonal strains in identical twins BF3 and BF4, which may have resulted from
517 maternal transfer. However, potential strain transmission between these infants living in
518 the same environment may also occur. Furthermore, wider studies involving both mothers
519 and twin babies (and other siblings) could provide details on the extent, timing and location
520 of transmission events between members of the same household.
521 Another aspect of comparative genomic analysis involved *in-silico* prediction of genes
522 belonging to GH families. This analysis revealed genome flexibility within *B. longum*, with
523 differences in GH family content between strains belonging to the same subspecies as
524 described previously; *B. infantis* predominantly enriched in GH families implicated in the
525 degradation of host-derived breast milk-associated dietary components like HMOs and *B.*
526 *longum* containing GH families involved in the metabolism of plant-derived substrates
527 (Milani et al., 2015a, Milani et al., 2016). Previously, Vatanen et al. (Vatanen et al., 2019)
528 suggested that the presence of the HMO gene cluster allowing for intracellular HMO
529 utilisation in *B. infantis* strains, in particular, confers a competitive advantage leading to
530 higher relative abundance of this subspecies in the early life microbiota. Our analysis of *B.*
531 *infantis* group identified the presence of glycosyl hydrolases associated with HMO
532 degradation in all isolates and revealed subspecies-specific differences in GH content
533 between pre- and post-weaning strains. Moreover, we observed differences in the number
534 of genes belonging to the most abundant GH families (e.g. GH43) between breast-fed and
535 formula-fed strains at different dietary stages, which can be linked to nutrient availability.
536 Surprisingly, we computationally and phenotypically identified closely related weaning and
537 post-weaning *B. longum* strains capable of metabolising HMOs (i.e. 2'-FL) in a formula-fed
538 baby that only received standard non-supplemented (i.e. no prebiotics or synthetic HMOs)

539 formula. The analysis of SNP variants in genes identified as glycosyl hydrolases predicted the
540 presence of missense mutations in both *B. longum* and *B. infantis* strains. Given that some
541 missense variants can compromise protein function (Miosge et al., 2015), our results
542 suggest potential functional differences that could further explain intra-strain and intra-
543 individual carbohydrate metabolism profiles of *B. longum*. However, experimental validation
544 would be essential to confirm the importance of variant predictions.
545 Recorded phenotypic data support the results of genomic analyses and further highlight
546 differences in carbohydrate utilisation profiles between and within *B. longum* and *B.*
547 *infantis*. As highlighted above, the ability of *B. infantis* to grow on different HMOs may
548 facilitate their early life establishment. Similarly, *B. longum* preference for plant-based
549 nutrients may be influencing their ability to persist within individual hosts through
550 significant dietary changes. Differential growth of strains that are genotypically similar on
551 various carbohydrate substrates and the ability of formula-fed strains to metabolise
552 selected HMOs suggest that *Bifidobacterium* possess an overall very broad repertoire of
553 genes for carbohydrate acquisition and metabolism that may be differentially switched on
554 and off in response to the presence of specific dietary components (Dworkin and Losick,
555 2001, Slager and Veening, 2016). Another explanation for these results may be a potential
556 influence of the intra-individual environment on epigenetic mechanisms in these bacteria.
557 One potential factor involved in this process may be a cooperative effort supported by
558 cross-feeding activities among *Bifidobacterium*, or between *Bifidobacterium* and other
559 members of the early life microbiota, e.g. *Bacteroides* and *Eubacterium* species (Rios-Covian
560 et al., 2013, Milani et al., 2015a, Schwab et al., 2017, Lawson et al., 2020). Indeed, the FISH
561 analysis revealed the presence of bacteria detected by probes Bac303 (bacteroides) and
562 ER482 (eubacterium) in faecal samples of both breast- and formula-fed infants, with intra-
563 individual variation at different dietary stages. Although *B. infantis* is principally known as a
564 specialist HMO-degrader, we did note growth of one of the *B. infantis* strains from formula-
565 fed baby FF1 on xylose. However, this growth profile was not consistent between
566 experiments, and therefore we did not pursue a fuller characterisation. However, future
567 examination of the ability of *B. longum* subsp. *infantis* to degrade a wider range of non-
568 HMO carbohydrate sources in early life could provide additional insight into carbohydrate
569 metabolism properties of this subspecies, and its role in ecosystem structuring during
570 transition to a more complex diet.

571 Glycan uptake analysis and proteomic investigation allowed us to determine mechanisms
572 which selected *B. longum* strains employ to metabolise different carbohydrates. A common
573 feature, based on the predicted activity of the most abundant proteins detected during
574 grown on the three substrates (cellobiose, LNnT and 2'-FL), was that they were all imported
575 and "selfishly" degraded intracellularly, therefore limiting release of degradation products
576 that could allow cross-feeding by other gut bacteria. This is in line with the carbohydrate
577 uptake analysis, where no peak for cellobiose, LNnT and 2'-FL degradation products could
578 be detected. Cellobiose uptake in B_25 occurs via a mechanism similar to *B. animalis* subsp.
579 *lactis* Bl-04 (*B. lactis*) (Andersen et al., 2013); cellobiose hydrolysis appears to be mediated
580 by the activity of three intracellular β-glucosidases, although further confirmatory
581 biochemical characterization of these enzyme is still required. B_25 was observed to utilize
582 LNnT using a pathway similar to that described in *B. longum* subsp. i*nfantis* whereby LNnT is
583 internalized via an ABC-transporter (B_25_00111-00113) followed by intracellular
584 degradation into constituent monosaccharides by a GH112 (B_25_00114) and an *N*-
585 acetylhexosamine 1-kinase (B_25_00115). LNnT degradation products are further
586 metabolized to fructose-6-phosphate by activities that include B_25_00116-00117
587 (galactose-1-phosphate urydyltranferase, UDP-glucose 4-epimerase, involved in the Leloir
588 pathway) and B_25_01030-01033 (for metabolism of *N*-acetylgalactosamine) prior to
589 entering the *Bifidobacterium* genus-specific fructose-6-phosphate phosphoketolase (F6PPK)
590 pathway (Ozcan and Sela, 2018). B_71 is predicted to deploy an ABC-transporter
591 (B_71_00974-00976) that allows uptake of intact 2'-FL that is subsequently hydrolysed to L-
592 fucose and lactose by the two predicted intracellular α-fucosidases GH29 (B_71_00982) and
593 GH95 (B_71_00983). L-fucose is further metabolized to L-lactate and pyruvate, via a
594 pathway of non-phosphorylated intermediates that include activities of L-fucose mutarotase
595 (B_71_00981), L-fucose dehydrogenase (B_71_00978), L-fuconate hydrolase (B_71_00977)
596 as previously described for *B. longum* subsp. *longum* SC596 (Garrido et al., 2016).
597 Considering that the proteins encoded by the aforementioned genes are located in the
598 cellobiose, LNnT and 2'-FL gene clusters that share high similarity and similar organization
599 with those found in equivalent systems in other *B. longum* and *B. lactis*, it is reasonable to
600 suggest that the gene clusters are related and may be the results of horizontal gene transfer
601 events between *B. longum*/*B. lactis* members residing in the infant gut microbiota.
602 Collectively, these data reflect inter- and intra-host phenotypic diversity of *B. longum* strains

603 in terms of their carbohydrate degradation capabilities and suggest that intra-individual
604 environment may influence epigenetic mechanisms in *Bifidobacterium*, resulting in
605 differential growth on carbohydrate substrates.
606 In conclusion, this research provides new insight into distinct genomic and phenotypic
607 abilities of *B. longum* species and strains isolated from the same individuals during the early
608 life developmental window by demonstrating that subspecies- and strain-specific
609 differences between members of *B. longum* sp. in infant hosts can be correlated to their
610 adaptation at specific age and diet stages.
611

612 **Limitations of the study**

613 Here, we used a combination of bioinformatic approaches and experimental techniques to
614 assess genomic and phenotypic abilities of *B. longum* species and strains isolated during the
615 early life developmental window. This study, however, is not without its limitations. One
616 important caveat is the small number of *B. infantis* strains (n=13) available for analysis, and
617 the fact that most of these strains (n=11) were isolated from a single breast-fed baby (BF2).
618 The examination of these strains provides important insight into the properties of *B. infantis*
619 during the transition from breastfeeding to more diversified diet, however it is difficult to
620 assess how representative these results are of wider population. In addition, only one strain
621 isolated from a formula-fed baby was identified as *B. infantis*, making it impossible to
622 examine properties of members of this subspecies within this dietary group and make
623 comparisons with breast-fed strains. Another important limitation is the fact that our strain
624 collection only contains one bacterial strain per time point. Inclusion of additional strains
625 could contribute further observations on inter-individual diversity of *Bifidobacterium* in
626 infant hosts and their functional properties. To examine bacterial communities in faecal
627 samples, we revisited and re-analysed the data generated using FISH, but this technique has
628 a detection limit (~10^6 bacterial cells (wet weight faeces)$^{-1}$) (Roger and McCartney, 2010).
629 Thus, FISH allows investigation of important bacterial groups, but faecal samples may
630 contain several organisms at levels below the methodological detection threshold. In
631 addition, this technique does not allow for tracking species-level changes. This limitation
632 could be addressed by the use of comprehensive sequencing methods, such as shotgun
633 metagenomics, combined with advanced computational methods to achieve strain-level
634 resolution. Furthermore, phenotypic investigation of carbohydrate metabolism properties

635 of *B. longum* revealed inconsistencies in growth of individual strains on certain
636 carbohydrates, including LNnT, cellobiose and mannose, and we therefore only explored
637 reproducible findings further with proteomics. Previously, variability in growth of *B. longum*
638 on mannose, even when analysing the same strain (*Bifidobacterium longum* NCC2705) has
639 been reported (Parche et al., 2007, Liu et al., 2011). Finally, no metadata on complementary
640 foods during weaning and infant diet post-weaning were available. This information could
641 allow bioinformatic predictions of carbohydrate degradation properties of *B. longum* to be
642 related to the specific dietary components present in weaning infant foods. Future
643 longitudinal studies could be designed to include these data.
644
645 **Resource Availability**
646
647 **Lead Contact**
648
649 Further information and requests for resources and reagents should be directed to and will
650 be fulfilled by the Lead Contact, Lindsay J. Hall (Lindsay.Hall@quadram.ac.uk).
651
652 **Materials Availability**
653
654 This study did not generate new unique reagents.
655
656 **Data and Code Availability**
657
658 The draft genomes of 75 *B. longum* isolates have been deposited to GOLD database at
659 https://img.jgi.doe.gov, GOLD Study ID: Gs0145337.
660 The proteomics data have been deposited to the ProteomeXchange Consortium
661 (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository with
662 dataset identifier PXD017277.
663
664 **Methods**
665

Journal Pre-proof

666 All methods can be found in the accompanying Transparent Methods supplemental file.

667

668 **Supplemental Information**

669

670 Supplemental Information can be found online at

671

672 **Acknowledgments**

673 We would like to thank Glycom A/S for the kind donation of purified HMOs: 2′-FL and LNnT.
674 The authors would also like to thank Prof Rob Kingsley and Mr Shabhonam Caim for
675 technical support and advice. This work was funded by a Wellcome Trust Investigator Award
676 (no. 100/974/C/13/Z); a BBSRC Norwich Research Park Bioscience Doctoral Training grant
677 no. BB/M011216/1 (supervisor LJH, student MK); an Institute Strategic Programme Gut
678 Microbes and Health grant no. BB/R012490/1 and its constituent projects
679 BBS/E/F/000PR10353 and BBS/E/F/000PR10356; and an Institute Strategic Programme Gut
680 Health and Food Safety grant no. BB/J004529/1 to LJH. LH was in receipt of a Medical
681 Research Council Intermediate Research Fellowship in Data Science (UK MED-BIO, grant no.
682 MR/L01632X/1). PBP and SLLR are grateful for support from The Research Council of
683 Norway (FRIPRO program, PBP: 250479), as well as the European Research Commission
684 Starting Grant Fellowship (awarded to PBP; 336355 - MicroDE). The funding bodies did not
685 contribute to the design of the study, collection, analysis, and interpretation of data or in
686 writing the manuscript.

687

688 **Author contributions**

689 LJH, LH, ALM and MK designed the overall study. ALM provided the unique *B. longum* strain
690 collection and extracted the DNA. ALM and LCR performed the FISH microbiota profiling
691 studies. MK prepared the DNA for WGS, performed all genomic analysis and visualisation, as
692 well as growth studies. MK also re-analysed the FISH data. SLLR, PBP, LJH and MK planned
693 metabolomics and proteomics studies. MK prepared samples for metabolomics and
694 proteomics. SLLR and MK performed the metabolomics and proteomics experiments and
695 SLLR analysed and visualised the resulting data. LJH and MK analysed the data, with input
696 and discussion from LH, and drafted the manuscript. SLLR, PBP, LH and ALM provided

697 providing further edits and co-writing of the final version. All authors read and approved the
698 final manuscript.
699

700 **Declaration of Interests**
701

702 The authors declare no competing interests.
703

704 **References**

705 AHMED, S., LUIS, A. S., BRAS, J. L., GHOSH, A., GAUTAM, S., GUPTA, M. N., FONTES, C. M. & GOYAL,
706 A. 2013. A novel alpha-L-arabinofuranosidase of family 43 glycoside hydrolase (Ct43Araf)
707 from Clostridium thermocellum. *PLoS One,* 8**,** e73575.
708 ANDERSEN, J. M., BARRANGOU, R., ABOU HACHEM, M., LAHTINEN, S. J., GOH, Y. J., SVENSSON, B. &
709 KLAENHAMMER, T. R. 2013. Transcriptional analysis of oligosaccharide utilization by
710 Bifidobacterium lactis Bl-04. *BMC Genomics,* 14**,** 312.
711 ARBOLEYA, S., BOTTACINI, F., O'CONNELL-MOTHERWAY, M., RYAN, C. A., ROSS, R. P., VAN
712 SINDEREN, D. & STANTON, C. 2018. Gene-trait matching across the Bifidobacterium longum
713 pan-genome reveals considerable diversity in carbohydrate catabolism among human infant
714 strains. *Bmc Genomics,* 19.
715 ASNICAR, F., MANARA, S., ZOLFO, M., TRUONG, D. T., SCHOLZ, M., ARMANINI, F., FERRETTI, P.,
716 GORFER, V., PEDROTTI, A., TETT, A. & SEGATA, N. 2017. Studying Vertical Microbiome
717 Transmission from Mothers to Infants by Strain-Level Metagenomic Profiling. *mSystems,* 2.
718 BACKHED, F., ROSWALL, J., PENG, Y., FENG, Q., JIA, H., KOVATCHEVA-DATCHARY, P., LI, Y., XIA, Y.,
719 XIE, H., ZHONG, H., KHAN, M. T., ZHANG, J., LI, J., XIAO, L., AL-AAMA, J., ZHANG, D., LEE, Y. S.,
720 KOTOWSKA, D., COLDING, C., TREMAROLI, V., YIN, Y., BERGMAN, S., XU, X., MADSEN, L.,
721 KRISTIANSEN, K., DAHLGREN, J. & WANG, J. 2015. Dynamics and Stabilization of the Human
722 Gut Microbiome during the First Year of Life. *Cell Host Microbe,* 17**,** 852.
723 BOKULICH, N. A., CHUNG, J., BATTAGLIA, T., HENDERSON, N., JAY, M., LI, H. L., LIEBER, A. D., WU, F.,
724 PEREZ-PEREZ, G. I., CHEN, Y., SCHWEIZER, W., ZHENG, X. H., CONTRERAS, M., DOMINGUEZ-
725 BELLO, M. G. & BLASER, M. J. 2016. Antibiotics, birth mode, and diet shape microbiome
726 maturation during early life. *Science Translational Medicine,* 8.
727 BROEKAERT, W. F., COURTIN, C. M., VERBEKE, K., VAN DE WIELE, T., VERSTRAETE, W. & DELCOUR, J.
728 A. 2011. Prebiotic and other health-related effects of cereal-derived arabinoxylans,
729 arabinoxylan-oligosaccharides, and xylooligosaccharides. *Crit Rev Food Sci Nutr,* 51**,** 178-94.
730 CHAPLIN, A. V., EFIMOV, B. A., SMEIANOV, V. V., KAFARSKAIA, L. I., PIKINA, A. P. & SHKOPOROV, A.
731 N. 2015. Intraspecies Genomic Diversity and Long-Term Persistence of Bifidobacterium
732 longum. *Plos One,* 10.
733 DAS, U. N. 2007. Breastfeeding prevents type 2 diabetes mellitus: but, how and why? *Am J Clin Nutr,*
734 85**,** 1436-7.
735 DE AGUERO, M. G., GANAL-VONARBURG, S. C., FUHRER, T., RUPP, S., UCHIMURA, Y., LI, H.,
736 STEINERT, A., HEIKENWALDER, M., HAPFELMEIER, S., SAUER, U., MCCOY, K. D. &
737 MACPHERSON, A. J. 2016. The maternal microbiota drives early postnatal innate immune
738 development. *Science,* 351**,** 1296-1301.
739 DOGRA, S., SAKWINSKA, O., SOH, S. E., NGOM-BRU, C., BRUCK, W. M., BERGER, B., BRUSSOW, H.,
740 LEE, Y. S., YAP, F., CHONG, Y. S., GODFREY, K. M., HOLBROOK, J. D. & GROUP, G. S. 2015.
741 Dynamics of infant gut microbiota are influenced by delivery mode and gestational duration
742 and are associated with subsequent adiposity. *mBio,* 6.

Journal Pre-proof

743 DWORKIN, J. & LOSICK, R. 2001. Linking nutritional status to gene activation and development.
744 *Genes Dev,* 15**,** 1051-4.
745 FENG, Q., LIANG, S. S., JIA, H. J., STADLMAYR, A., TANG, L. Q., LAN, Z., ZHANG, D. Y., XIA, H. H., XU, X.
746 Y., JIE, Z. Y., SU, L. L., LI, X. P., LI, X., LI, J. H., XIAO, L., HUBER-SCHONAUER, U., NIEDERSEER,
747 D., XU, X., AL-AAMA, J. Y., YANG, H. M., WANG, J., KRISTIANSEN, K., ARUMUGAM, M., TILG,
748 H., DATZ, C. & WANG, J. 2015. Gut microbiome development along the colorectal adenoma-
749 carcinoma sequence. *Nature Communications,* 6.
750 FLORINDO, R. N., SOUZA, V. P., MANZINE, L. R., CAMILO, C. M., MARANA, S. R., POLIKARPOV, I. &
751 NASCIMENTO, A. S. 2018. Structural and biochemical characterization of a GH3 beta-
752 glucosidase from the probiotic bacteria Bifidobacterium adolescentis. *Biochimie,* 148**,** 107-
753 115.
754 FORBES, J. D., AZAD, M. B., VEHLING, L., TUN, H. M., KONYA, T. B., GUTTMAN, D. S., FIELD, C. J.,
755 LEFEBVRE, D., SEARS, M. R., BECKER, A. B., MANDHANE, P. J., TURVEY, S. E., MORAES, T. J.,
756 SUBBARAO, P., SCOTT, J. A., KOZYRSKYJ, A. L. & INVESTIGATORS, C. S. 2018. Association of
757 Exposure to Formula in the Hospital and Subsequent Infant Feeding Practices With Gut
758 Microbiota and Risk of Overweight in the First Year of Life. *Jama Pediatrics,* 172.
759 GARRIDO, D., BARILE, D. & MILLS, D. A. 2012. A molecular basis for bifidobacterial enrichment in the
760 infant gastrointestinal tract. *Adv Nutr,* 3**,** 415S-21S.
761 GARRIDO, D., RUIZ-MOYANO, S., KIRMIZ, N., DAVIS, J. C., TOTTEN, S. M., LEMAY, D. G., UGALDE, J. A.,
762 GERMAN, J. B., LEBRILLA, C. B. & MILLS, D. A. 2016. A novel gene cluster allows preferential
763 utilization of fucosylated milk oligosaccharides in Bifidobacterium longum subsp longum
764 SC596. *Scientific Reports,* 6.
765 GARRIDO, D., RUIZ-MOYANO, S., LEMAY, D. G., SELA, D. A., GERMAN, J. B. & MILLS, D. A. 2015.
766 Comparative transcriptomics reveals key differences in the response to milk oligosaccharides
767 of infant gut-associated bifidobacteria. *Sci Rep,* 5**,** 13517.
768 HALD, S., SCHIOLDAN, A. G., MOORE, M. E., DIGE, A., LAERKE, H. N., AGNHOLT, J., KNUDSEN, K. E. B.,
769 HERMANSEN, K., MARCO, M. L., GREGERSEN, S. & DAHLERUP, J. F. 2016. Effects of
770 Arabinoxylan and Resistant Starch on Intestinal Microbiota and Short-Chain Fatty Acids in
771 Subjects with Metabolic Syndrome: A Randomised Crossover Study. *Plos One,* 11.
772 HEIKKILA, M. P. & SARIS, P. E. 2003. Inhibition of Staphylococcus aureus by the commensal bacteria
773 of human milk. *J Appl Microbiol,* 95**,** 471-8.
774 HINZ, S. W., PASTINK, M. I., VAN DEN BROEK, L. A., VINCKEN, J. P. & VORAGEN, A. G. 2005.
775 Bifidobacterium longum endogalactanase liberates galactotriose from type I galactans. *Appl*
776 *Environ Microbiol,* 71**,** 5501-10.
777 ICHINOSE, H., YOSHIDA, M., FUJIMOTO, Z. & KANEKO, S. 2008. Characterization of a modular enzyme
778 of exo-1,5-alpha-L-arabinofuranosidase and arabinan binding module from Streptomyces
779 avermitilis NBRC14893. *Appl Microbiol Biotechnol,* 80**,** 399-408.
780 IP, S., CHUNG, M., RAMAN, G., CHEW, P., MAGULA, N., DEVINE, D., TRIKALINOS, T. & LAU, J. 2007.
781 Breastfeeding and maternal and infant health outcomes in developed countries. *Evid Rep*
782 *Technol Assess (Full Rep)***,** 1-186.
783 JAMES, K., MOTHERWAY, M. O., BOTTACINI, F. & VAN SINDEREN, D. 2016. Bifidobacterium breve
784 UCC2003 metabolises the human milk oligosaccharides lacto-N-tetraose and lacto-N-neo-
785 tetraose through overlapping, yet distinct pathways. *Sci Rep,* 6**,** 38560.
786 KATAYAMA, T. 2016. Host-derived glycans serve as selected nutrients for the gut microbe: human
787 milk oligosaccharides and bifidobacteria. *Bioscience Biotechnology and Biochemistry,* 80**,**
788 621-632.
789 KITAOKA, M. 2012. Bifidobacterial enzymes involved in the metabolism of human milk
790 oligosaccharides. *Adv Nutr,* 3**,** 422S-9S.
791 KOENIG, J. E., SPOR, A., SCALFONE, N., FRICKER, A. D., STOMBAUGH, J., KNIGHT, R., ANGENENT, L. T.
792 & LEY, R. E. 2011. Succession of microbial consortia in the developing infant gut microbiome.
793 *Proc Natl Acad Sci U S A,* 108 Suppl 1**,** 4578-85.

794 LAWSON, M. A. E., O'NEILL, I. J., KUJAWSKA, M., GOWRINADH JAVVADI, S., WIJEYESEKERA, A.,
795 FLEGG, Z., CHALKLEN, L. & HALL, L. J. 2020. Breast milk-derived human milk oligosaccharides
796 promote Bifidobacterium interactions within a single ecosystem. *ISME J,* 14**,** 635-648.
797 LAX, S., SMITH, D. P., HAMPTON-MARCELL, J., OWENS, S. M., HANDLEY, K. M., SCOTT, N. M.,
798 GIBBONS, S. M., LARSEN, P., SHOGAN, B. D., WEISS, S., METCALF, J. L., URSELL, L. K.,
799 VAZQUEZ-BAEZA, Y., VAN TREUREN, W., HASAN, N. A., GIBSON, M. K., COLWELL, R., DANTAS,
800 G., KNIGHT, R. & GILBERT, J. A. 2014. Longitudinal analysis of microbial interaction between
801 humans and the indoor environment. *Science,* 345**,** 1048-52.
802 LIU, D., WANG, S., XU, B., GUO, Y., ZHAO, J., LIU, W., SUN, Z., SHAO, C., WEI, X., JIANG, Z., WANG, X.,
803 LIU, F., WANG, J., HUANG, L., HU, D., HE, X., RIEDEL, C. U. & YUAN, J. 2011. Proteomics
804 analysis of Bifidobacterium longum NCC2705 growing on glucose, fructose, mannose, xylose,
805 ribose, and galactose. *Proteomics,* 11**,** 2628-38.
806 LUIS, A. S., BRIGGS, J., ZHANG, X., FARNELL, B., NDEH, D., LABOUREL, A., BASLE, A., CARTMELL, A.,
807 TERRAPON, N., STOTT, K., LOWE, E. C., MCLEAN, R., SHEARER, K., SCHUCKEL, J., VENDITTO, I.,
808 RALET, M. C., HENRISSAT, B., MARTENS, E. C., MOSIMANN, S. C., ABBOTT, D. W. & GILBERT,
809 H. J. 2018. Dietary pectic glycans are degraded by coordinated enzyme pathways in human
810 colonic Bacteroides. *Nat Microbiol,* 3**,** 210-219.
811 MAGNE, F., HACHELAF, W., SUAU, A., BOUDRAA, G., MANGIN, I., TOUHAMI, M., BOUZIANE-NEDJADI,
812 K. & POCHART, P. 2006. A longitudinal study of infant faecal microbiota during weaning.
813 *FEMS Microbiol Ecol,* 58**,** 563-71.
814 MAKINO, H., KUSHIRO, A., ISHIKAWA, E., KUBOTA, H., GAWAD, A., SAKAI, T., OISHI, K., MARTIN, R.,
815 BEN-AMOR, K., KNOL, J. & TANAKA, R. 2013. Mother-to-infant transmission of intestinal
816 bifidobacterial strains has an impact on the early development of vaginally delivered infant's
817 microbiota. *PLoS One,* 8**,** e78331.
818 MAKINO, H., KUSHIRO, A., ISHIKAWA, E., MUYLAERT, D., KUBOTA, H., SAKAI, T., OISHI, K., MARTIN,
819 R., BEN AMOR, K., OOZEER, R., KNOL, J. & TANAKA, R. 2011. Transmission of intestinal
820 Bifidobacterium longum subsp. longum strains from mother to infant, determined by
821 multilocus sequencing typing and amplified fragment length polymorphism. *Appl Environ*
822 *Microbiol,* 77**,** 6788-93.
823 MALDONADO-GOMEZ, M. X., MARTINEZ, I., BOTTACINI, F., O'CALLAGHAN, A., VENTURA, M., VAN
824 SINDEREN, D., HILLMANN, B., VANGAY, P., KNIGHTS, D., HUTKINS, R. W. & WALTER, J. 2016.
825 Stable Engraftment of Bifidobacterium longum AH1206 in the Human Gut Depends on
826 Individualized Features of the Resident Microbiome. *Cell Host Microbe,* 20**,** 515-526.
827 MARCOBAL, A. & SONNENBURG, J. L. 2012. Human milk oligosaccharide consumption by intestinal
828 microbiota. *Clinical Microbiology and Infection,* 18**,** 12-15.
829 MARTIN, R., MAKINO, H., CETINYUREK YAVUZ, A., BEN-AMOR, K., ROELOFS, M., ISHIKAWA, E.,
830 KUBOTA, H., SWINKELS, S., SAKAI, T., OISHI, K., KUSHIRO, A. & KNOL, J. 2016. Early-Life
831 Events, Including Mode of Delivery and Type of Feeding, Siblings and Gender, Shape the
832 Developing Gut Microbiota. *PLoS One,* 11**,** e0158498.
833 MATTARELLI, P., BONAPARTE, C., POT, B. & BIAVATI, B. 2008. Proposal to reclassify the three
834 biotypes of Bifidobacterium longum as three subspecies: Bifidobacterium longum subsp.
835 longum subsp. nov., Bifidobacterium longum subsp. infantis comb. nov. and Bifidobacterium
836 longum subsp. suis comb. nov. *Int J Syst Evol Microbiol,* 58**,** 767-72.
837 MCKEEN, S., YOUNG, W., MULLANEY, J., FRASER, K., MCNABB, W. C. & ROY, N. C. 2019. Infant
838 Complementary Feeding of Prebiotics for the Microbiome and Immunity. *Nutrients,* 11.
839 MIKAMI, K., KIMURA, M. & TAKAHASHI, H. 2012. Influence of maternal bifidobacteria on the
840 development of gut bifidobacteria in infants. *Pharmaceuticals (Basel),* 5**,** 629-42.
841 MILANI, C., LUGLI, G. A., DURANTI, S., TURRONI, F., MANCABELLI, L., FERRARIO, C., MANGIFESTA, M.,
842 HEVIA, A., VIAPPIANI, A., SCHOLZ, M., ARIOLI, S., SANCHEZ, B., LANE, J., WARD, D. V., HICKEY,
843 R., MORA, D., SEGATA, N., MARGOLLES, A., VAN SINDEREN, D. & VENTURA, M. 2015a.

844 Bifidobacteria exhibit social behavior through carbohydrate resource sharing in the gut.
845 *Scientific Reports,* 5.
846 MILANI, C., MANCABELLI, L., LUGLI, G. A., DURANTI, S., TURRONI, F., FERRARIO, C., MANGIFESTA, M.,
847 VIAPPIANI, A., FERRETTI, P., GORFER, V., TETT, A., SEGATA, N., VAN SINDEREN, D. &
848 VENTURA, M. 2015b. Exploring Vertical Transmission of Bifidobacteria from Mother to Child.
849 *Appl Environ Microbiol,* 81**,** 7078-87.
850 MILANI, C., TURRONI, F., DURANTI, S., LUGLI, G. A., MANCABELLI, L., FERRARIO, C., VAN SINDEREN,
851 D. & VENTURA, M. 2016. Genomics of the Genus Bifidobacterium Reveals Species-Specific
852 Adaptation to the Glycan-Rich Gut Environment. *Applied and Environmental Microbiology,*
853 82**,** 980-991.
854 MILLS, S., STANTON, C., LANE, J. A., SMITH, G. J. & ROSS, R. P. 2019. Precision Nutrition and the
855 Microbiome, Part I: Current State of the Science. *Nutrients,* 11.
856 MIOSGE, L. A., FIELD, M. A., SONTANI, Y., CHO, V., JOHNSON, S., PALKOVA, A., BALAKISHNAN, B.,
857 LIANG, R., ZHANG, Y., LYON, S., BEUTLER, B., WHITTLE, B., BERTRAM, E. M., ENDERS, A.,
858 GOODNOW, C. C. & ANDREWS, T. D. 2015. Comparison of predicted and actual
859 consequences of missense mutations. *Proc Natl Acad Sci U S A,* 112**,** E5189-98.
860 O'SULLIVAN, A., FARVER, M. & SMILOWITZ, J. T. 2015. The Influence of Early Infant-Feeding Practices
861 on the Intestinal Microbiome and Body Composition in Infants. *Nutr Metab Insights,* 8**,** 1-9.
862 ODAMAKI, T., BOTTACINI, F., KATO, K., MITSUYAMA, E., YOSHIDA, K., HORIGOME, A., XIAO, J. Z. &
863 VAN SINDEREN, D. 2018. Genomic diversity and distribution of Bifidobacterium longum
864 subsp longum across the human lifespan. *Scientific Reports,* 8.
865 OKI, K., AKIYAMA, T., MATSUDA, K., GAWAD, A., MAKINO, H., ISHIKAWA, E., OISHI, K., KUSHIRO, A. &
866 FUJIMOTO, J. 2018. Long-term colonization exceeding six years from early infancy of
867 Bifidobacterium longum subsp. longum in human gut. *Bmc Microbiology,* 18.
868 OLSZAK, T., AN, D. D., ZEISSIG, S., VERA, M. P., RICHTER, J., FRANKE, A., GLICKMAN, J. N., SIEBERT, R.,
869 BARON, R. M., KASPER, D. L. & BLUMBERG, R. S. 2012. Microbial Exposure During Early Life
870 Has Persistent Effects on Natural Killer T Cell Function. *Science,* 336**,** 489-493.
871 ORTEGA-GARCIA, J. A., KLOOSTERMAN, N., ALVAREZ, L., TOBARRA-SANCHEZ, E., CARCELES-ALVAREZ,
872 A., PASTOR-VALERO, R., LOPEZ-HERNANDEZ, F. A., SANCHEZ-SOLIS, M. & CLAUDIO, L. 2018.
873 Full Breastfeeding and Obesity in Children: A Prospective Study from Birth to 6 Years. *Child*
874 *Obes,* 14**,** 327-337.
875 OZCAN, E. & SELA, D. A. 2018. Inefficient Metabolism of the Human Milk Oligosaccharides Lacto-N-
876 tetraose and Lacto-N-neotetraose Shifts Bifidobacterium longum subsp. infantis Physiology.
877 *Frontiers in Nutrition,* 5.
878 PALMER, C., BIK, E. M., DIGIULIO, D. B., RELMAN, D. A. & BROWN, P. O. 2007. Development of the
879 human infant intestinal microbiota. *PLoS Biol,* 5**,** e177.
880 PARCHE, S., AMON, J., JANKOVIC, I., REZZONICO, E., BELEUT, M., BARUTCU, H., SCHENDEL, I., EDDY,
881 M. P., BURKOVSKI, A., ARIGONI, F. & TITGEMEYER, F. 2007. Sugar transport systems of
882 Bifidobacterium longum NCC2705. *J Mol Microbiol Biotechnol,* 12**,** 9-19.
883 POKUSAEVA, K., FITZGERALD, G. F. & VAN SINDEREN, D. 2011. Carbohydrate metabolism in
884 Bifidobacteria. *Genes and Nutrition,* 6**,** 285-306.
885 POZZO, T., PASTEN, J. L., KARLSSON, E. N. & LOGAN, D. T. 2010. Structural and functional analyses of
886 beta-glucosidase 3B from Thermotoga neapolitana: a thermostable three-domain
887 representative of glycoside hydrolase 3. *J Mol Biol,* 397**,** 724-39.
888 RENZ, H., BRANDTZAEG, P. & HORNEF, M. 2012. The impact of perinatal immune development on
889 mucosal homeostasis and chronic inflammation. *Nature Reviews Immunology,* 12**,** 9-23.
890 RIOS-COVIAN, D., ARBOLEYA, S., HERNANDEZ-BARRANCO, A. M., ALVAREZ-BUYLLA, J. R., RUAS-
891 MADIEDO, P., GUEIMONDE, M. & DE LOS REYES-GAVILAN, C. G. 2013. Interactions between
892 Bifidobacterium and Bacteroides species in cofermentations are affected by carbon sources,
893 including exopolysaccharides produced by bifidobacteria. *Appl Environ Microbiol,* 79**,** 7518-
894 24.

Journal Pre-proof

Journal Pre-proof

895 RIVIERE, A., SELAK, M., LANTIN, D., LEROY, F. & DE VUYST, L. 2016. Bifidobacteria and Butyrate-
896 Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human
897 Gut. *Front Microbiol,* 7**,** 979.
898 ROBERFROID, M. B. 2007. Inulin-type fructans: functional food ingredients. *J Nutr,* 137**,** 2493S-2502S.
899 ROGER, L. C., COSTABILE, A., HOLLAND, D. T., HOYLES, L. & MCCARTNEY, A. L. 2010. Examination of
900 faecal Bifidobacterium populations in breast- and formula-fed infants during the first 18
901 months of life. *Microbiology-Sgm,* 156**,** 3329-3341.
902 ROGER, L. C. & MCCARTNEY, A. L. 2010. Longitudinal investigation of the faecal microbiota of healthy
903 full-term infants using fluorescence in situ hybridization and denaturing gradient gel
904 electrophoresis. *Microbiology-Sgm,* 156**,** 3317-3328.
905 SCHWAB, C., RUSCHEWEYH, H. J., BUNESOVA, V., PHAM, V. T., BEERENWINKEL, N. & LACROIX, C.
906 2017. Trophic Interactions of Infant Bifidobacteria and Eubacterium hallii during L-Fucose
907 and Fucosyllactose Degradation. *Frontiers in Microbiology,* 8.
908 SELA, D. A., CHAPMAN, J., ADEUYA, A., KIM, J. H., CHEN, F., WHITEHEAD, T. R., LAPIDUS, A.,
909 ROKHSAR, D. S., LEBRILLA, C. B., GERMAN, J. B., PRICE, N. P., RICHARDSON, P. M. & MILLS, D.
910 A. 2008. The genome sequence of Bifidobacterium longum subsp infantis reveals
911 adaptations for milk utilization within the infant microbiome. *Proceedings of the National*
912 *Academy of Sciences of the United States of America,* 105**,** 18964-18969.
913 SELA, D. A., GARRIDO, D., LERNO, L., WU, S., TAN, K., EOM, H. J., JOACHIMIAK, A., LEBRILLA, C. B. &
914 MILLS, D. A. 2012. Bifidobacterium longum subsp. infantis ATCC 15697 alpha-fucosidases are
915 active on fucosylated human milk oligosaccharides. *Appl Environ Microbiol,* 78**,** 795-803.
916 SELA, D. A. & MILLS, D. A. 2010. Nursing our microbiota: molecular linkages between bifidobacteria
917 and milk oligosaccharides. *Trends in Microbiology,* 18**,** 298-307.
918 SHAO, Y., FORSTER, S. C., TSALIKI, E., VERVIER, K., STRANG, A., SIMPSON, N., KUMAR, N., STARES, M.
919 D., RODGER, A., BROCKLEHURST, P., FIELD, N. & LAWLEY, T. D. 2019. Stunted microbiota and
920 opportunistic pathogen colonization in caesarean-section birth. *Nature,* 574**,** 117-121.
921 SIVAN, A., CORRALES, L., HUBERT, N., WILLIAMS, J. B., AQUINO-MICHAELS, K., EARLEY, Z. M.,
922 BENYAMIN, F. W., LEI, Y. M., JABRI, B., ALEGRE, M. L., CHANG, E. B. & GAJEWSKI, T. F. 2015.
923 Commensal Bifidobacterium promotes antitumor immunity and facilitates anti-PD-L1
924 efficacy. *Science,* 350**,** 1084-1089.
925 SLAGER, J. & VEENING, J. W. 2016. Hard-Wired Control of Bacterial Processes by Chromosomal Gene
926 Location. *Trends Microbiol,* 24**,** 788-800.
927 STIEMSMA, L. T. & MICHELS, K. B. 2018. The Role of the Microbiome in the Developmental Origins of
928 Health and Disease. *Pediatrics,* 141.
929 SUZUKI, H., MURAKAMI, A. & YOSHIDA, K. 2013. Motif-guided identification of a glycoside hydrolase
930 family 1 alpha-L-arabinofuranosidase in Bifidobacterium adolescentis. *Biosci Biotechnol*
931 *Biochem,* 77**,** 1709-14.
932 TANG, W. H. W., KITAI, T. & HAZEN, S. L. 2017. Gut Microbiota in Cardiovascular Health and Disease.
933 *Circulation Research,* 120**,** 1183-1196.
934 THONGARAM, T., HOEFLINGER, J. L., CHOW, J. & MILLER, M. J. 2017. Human milk oligosaccharide
935 consumption by probiotic and human-associated bifidobacteria and lactobacilli. *J Dairy Sci,*
936 100**,** 7825-7833.
937 TURNBAUGH, P. J., LEY, R. E., MAHOWALD, M. A., MAGRINI, V., MARDIS, E. R. & GORDON, J. I. 2006.
938 An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature,*
939 444**,** 1027-1031.
940 TURRONI, F., FORONI, E., PIZZETTI, P., GIUBELLINI, V., RIBBERA, A., MERUSI, P., CAGNASSO, P.,
941 BIZZARRI, B., DE'ANGELIS, G. L., SHANAHAN, F., VAN SINDEREN, D. & VENTURA, M. 2009.
942 Exploring the Diversity of the Bifidobacterial Population in the Human Intestinal Tract.
943 *Applied and Environmental Microbiology,* 75**,** 1534-1545.
944 TURRONI, F., PEANO, C., PASS, D. A., FORONI, E., SEVERGNINI, M., CLAESSON, M. J., KERR, C.,
945 HOURIHANE, J., MURRAY, D., FULIGNI, F., GUEIMONDE, M., MARGOLLES, A., DE BELLIS, G.,

Journal Pre-proof

946 O'TOOLE, P. W., VAN SINDEREN, D., MARCHESI, J. R. & VENTURA, M. 2012. Diversity of
947 Bifidobacteria within the Infant Gut Microbiota. *Plos One,* 7.
948 VATANEN, T., PLICHTA, D. R., SOMANI, J., MUNCH, P. C., ARTHUR, T. D., HALL, A. B., RUDOLF, S.,
949 OAKELEY, E. J., KE, X., YOUNG, R. A., HAISER, H. J., KOLDE, R., YASSOUR, M., LUOPAJARVI, K.,
950 SILJANDER, H., VIRTANEN, S. M., ILONEN, J., UIBO, R., TILLMANN, V., MOKUROV, S.,
951 DORSHAKOVA, N., PORTER, J. A., MCHARDY, A. C., LAHDESMAKI, H., VLAMAKIS, H.,
952 HUTTENHOWER, C., KNIP, M. & XAVIER, R. J. 2019. Genomic variation and strain-specific
953 functional adaptation in the human gut microbiome during early life. *Nat Microbiol,* 4**,** 470-
954 479.
955 VENTURA, M., O'FLAHERTY, S., CLAESSON, M. J., TURRONI, F., KLAENHAMMER, T. R., VAN SINDEREN,
956 D. & O'TOOLE, P. W. 2009. Genome-scale analyses of health-promoting bacteria:
957 probiogenomics. *Nat Rev Microbiol,* 7**,** 61-71.
958 VIBORG, A. H., SORENSEN, K. I., GILAD, O., STEEN-JENSEN, D. B., DILOKPIMOL, A., JACOBSEN, S. &
959 SVENSSON, B. 2013. Biochemical and kinetic characterisation of a novel xylooligosaccharide-
960 upregulated GH43 beta-D-xylosidase/alpha-L-arabinofuranosidase (BXA43) from the
961 probiotic Bifidobacterium animalis subsp lactis BB-12. *Amb Express,* 3.
962 WAMPACH, L., HEINTZ-BUSCHART, A., HOGAN, A., MULLER, E. E. L., NARAYANASAMY, S., LACZNY, C.
963 C., HUGERTH, L. W., BINDL, L., BOTTU, J., ANDERSSON, A. F., DE BEAUFORT, C. & WILMES, P.
964 2017. Colonization and Succession within the Human Gut Microbiome by Archaea, Bacteria,
965 and Microeukaryotes during the First Year of Life. *Front Microbiol,* 8**,** 738.
966 WATSON, D., MOTHERWAY, M. O., SCHOTERMAN, M. H. C., VAN NEERVEN, R. J. J., NAUTA, A. & VAN
967 SINDEREN, D. 2013. Selective carbohydrate utilization by lactobacilli and bifidobacteria.
968 *Journal of Applied Microbiology,* 114**,** 1132-1146.
969 YAMADA, C., GOTOH, A., SAKANAKA, M., HATTIE, M., STUBBS, K. A., KATAYAMA-IKEGAMI, A.,
970 HIROSE, J., KURIHARA, S., ARAKAWA, T., KITAOKA, M., OKUDA, S., KATAYAMA, T. &
971 FUSHINOBU, S. 2017. Molecular Insight into Evolution of Symbiosis between Breast-Fed
972 Infants and a Member of the Human Gut Microbiome Bifidobacterium longum. *Cell Chemical*
973 *Biology,* 24**,** 515-+.
974 YAN, S., WEI, P. C., CHEN, Q., CHEN, X., WANG, S. C., LI, J. R. & GAO, C. 2018. Functional and
975 structural characterization of a beta-glucosidase involved in saponin metabolism from
976 intestinal bacteria. *Biochem Biophys Res Commun,* 496**,** 1349-1356.
977 YANOKURA, E., OKI, K., MAKINO, H., MODESTO, M., POT, B., MATTARELLI, P., BIAVATI, B. &
978 WATANABE, K. 2015. Subspeciation of Bifidobacterium longum by multilocus approaches
979 and amplified fragment length polymorphism: Description of B. longum subsp. suillum
980 subsp. nov., isolated from the faeces of piglets. *Syst Appl Microbiol,* 38**,** 305-14.
981

982 **Figure and Scheme Legends**

983

984 **Figure 1.** Proportional representation of bacterial populations in the faecal microbiota of a)
985 breast-fed and b) formula-fed infants based on FISH analysis. Numbers are expressed as
986 percentage of the total bacterial population obtained using DAPI. The vertical solid black
987 lines mark the different dietary phases in each infant (pre-weaning, weaning and post-
988 weaning). Oligonucleotide probes used to determine bacterial populations: Bif164 -
989 most *Bifidobacterium* species and *Parascardovia denticolens*, Bac303 - most members of the
990 genus *Bacteroides*, some *Parabacteroides* and *Prevotella* species, *Paraprevotella*,

Journal Pre-proof

991 *Xylanibacter*, *Barnesiella* species and *Odoribacter splanchnicus*, ER482 - most members
992 of Clostridium cluster XIVa, Ato291 - *Cryptobacterium curtum*, *Gordonibacter*
993 *pamelaeae*, *Paraeggerthella hongkongensis*, all *Eggerthella*, *Collinsella*,
994 *Olsenella* and *Atopobium* species, Chis150 - most members of *Clostridium* cluster I, all
995 members of *Clostridium* cluster II, EC1531 - *Escherichia coli*, Lab158 - all *Oenococcus*,
996 *Vagococcus*, *Melissococcus*, *Tetragenococcus*, *Enterococcus*, *Catellicoccus*, *Paralactobacillus*,
997 *Pediococcus* and *Lactococcus* species, most *Lactobacillus*,
998 *Weissella* and *Leuconostoc* species. See also Table S1.
999
1000 **Figure 2.** Identification and relatedness of *B. longum* strains. a) Sampling scheme and strain
1001 identification within individual breast-fed (BF1-BF5) and formula-fed (FF1-FF3 and FF5)
1002 infants based on average nucleotide identity values (ANI). The three levels of shading mark
1003 different dietary phases: pre-weaning, weaning, and post-weaning. b) Relatedness of *B.*
1004 *longum* strains based on core proteins. Coloured strips represent isolation period (pre-
1005 weaning, weaning and post-weaning) and isolation source (individual infants), respectively.
1006 See also Table2, Table S3 and Table S4.
1007
1008 **Figure 3.** Pairwise SNP distances between *B. longum* strains of the same subspecies within
1009 individual infants. Individual points show data distribution, diamonds indicate the group
1010 mean, box plots show group median and interquartile range. See also Table S6.
1011
1012 **Figure 4.** Gene-loss events and abundance of GH families within *B. longum* subspecies. Pie
1013 charts superimposed on the whole genome SNP tree represent predicted GH family gain-
1014 loss events within *B. longum* and *B. infantis* lineages. Due to the size of the tree, examples
1015 of detailed gain loss events have been provided for main lineages, as well as baby BF2
1016 (strains highlighted with light blue) and BF5 (strains highlighted with light purple). Heatmap
1017 represents abundance of specific GH families predicted in analysed *B. longum* strains. See
1018 also Table S8 and Table S9.
1019
1020 **Figure 5.** Growth performance of *B. longum* strains on different carbon sources. Heatmap
1021 displays the difference in average growth of triplicates between T_2 (30 min) and T_{end} (48
1022 hours). Moderate growth is considered above 0.15 difference in OD from time T_2, high

1023 growth above 0.25 difference in OD from time T_2. Asterisks represent strains for which
1024 inconsistent growth was recorded (difference in OD of at least 0.15 between any of the
1025 duplicates in the triplicate experiment). See also Table S12.
1026
1027 **Figure 6.** HPAEC-PAD traces showing mono-, di- and oligo-saccharides detected in the
1028 supernatant of either B_25 or B_71 single cultures during growth in mMRS supplemented
1029 with (a) cellobiose; (b) LNnT; (c) 2'-FL. The data are representative of biological triplicates.
1030 Abbreviations: LNnT, Lacto-*N*-neotetraose; Glc, glucose; Glc2, cellobiose; 2'-FL, 2'-
1031 fucosyllactose. Panel on the right shows (a) cellobiose; (b) LNnT; (c) 2'-FL utilization clusters
1032 in B_25 and B_71 and proteomic detection of the corresponding proteins during growth on
1033 HMOs. Heat maps above genes show the LFQ detection levels for the corresponding
1034 proteins in triplicates grown on glucose (G); cellobiose (C); LNnT (L); and 2'-FL (F). Numbers
1035 between genes indicate percent identity between corresponding genes in homologous PULs
1036 relative to strains B_25 and B_71. Numbers below each gene show the locus tag in the
1037 corresponding genome. Locus tag numbers are abbreviated with the last numbers after the
1038 second hyphen (for example B_25_XXXXX). The locus tag prefix for each strain is indicated in
1039 parenthesis beside the organism name. See also Table S13 and Table S14.
1040
1041 **<u>Supplemental Table titles</u>**
1042
1043 **Table S1.** FISH probes used to enumerate bacterial groups in infants (adapted from Roger &
1044 McCartney (2010)), proportional representation of FISH data for each baby, summative
1045 analysis of bacterial populations in relation to dietary phases (separate tabs). Related to
1046 Figure 1.
1047
1048 **Table S2.** Demographic profiles of infants in this study, genomic data on the isolates in this
1049 study, list of publicly available genomes used in this study (separate tabs). Related to Figure
1050 2.
1051
1052 **Table S3.** Average nucleotide identity matrix for the isolates used in this study. Related to
1053 Figure 2.
1054

Journal Pre-proof

Journal Pre-proof

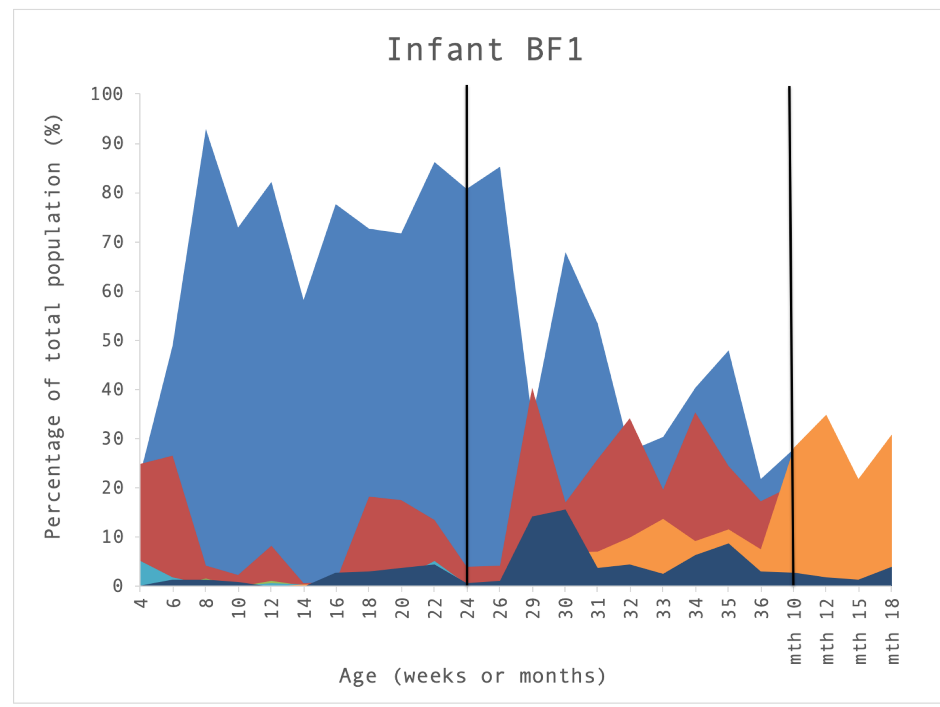
1055 **Table S4.** Gene presence-absence matrix for the *B. longum* isolates used in this study.
1056 Related to Figure 2.
1057
1058 **Table S5.** List of overrepresented genes associated with breast- and formula-feeding
1059 detected using Scoary (separate tabs). Related to Figure 2.
1060
1061 **Table S6.** Pairwise SNP distance matrix between isolates within individual infants. Related to
1062 Figure 3.
1063
1064 **Table S7.** Summary of COG categories identified in *B. longum* isolates. COG categories
1065 legend, COG categories identified in all *B. longum* samples, as well as in specific groups: *B.*
1066 *infantis*, *B. longum*, isolates from breast-fed infants, isolates from formula-fed infants
1067 (separate tabs). Related to Figure 4.
1068
1069 **Table S8.** Summary of GH families identified in all *B. longum* isolates, as well as in specific
1070 groups: *B. longum* across all dietary phases, *B. longum* pre-weaning, *B. longum* post-
1071 weaning, *B. longum* isolates from breast-fed infants, *B. longum* isolates from formula-fed
1072 infants, comparison between *B. longum* breast- and formula-fed groups, *B. infantis*, *B.*
1073 *infantis* pre-weaning, *B. infantis* post-weaning (separate tabs). Related to Figure 4.
1074
1075 **Table S9.** Summary of enzyme classes identified in all *B. longum* isolates, as well as in
1076 specific groups: *B. longum* across all dietary phases, *B. longum* pre-weaning, *B. longum* post-
1077 weaning and *B. infantis* (separate tabs). Related to Figure 4.
1078
1079 **Table S10.** Prediction of glycosyl hydrolase family gain-loss events in *B. longum* isolates.
1080 Related to Figure 4.
1081
1082 **Table S11.** Summary of SNPs in predicted glycosyl hydrolases in breast-fed and formula-fed
1083 babies. Related to Figure 3 and Figure 4.
1084
1085 **Table S12.** Results of growth experiments performed for selected *B. longum* and *B. infantis*
1086 strains from nine babies. Strains (1%v/v) were grown in modified (m)MRS (pH 6.8)

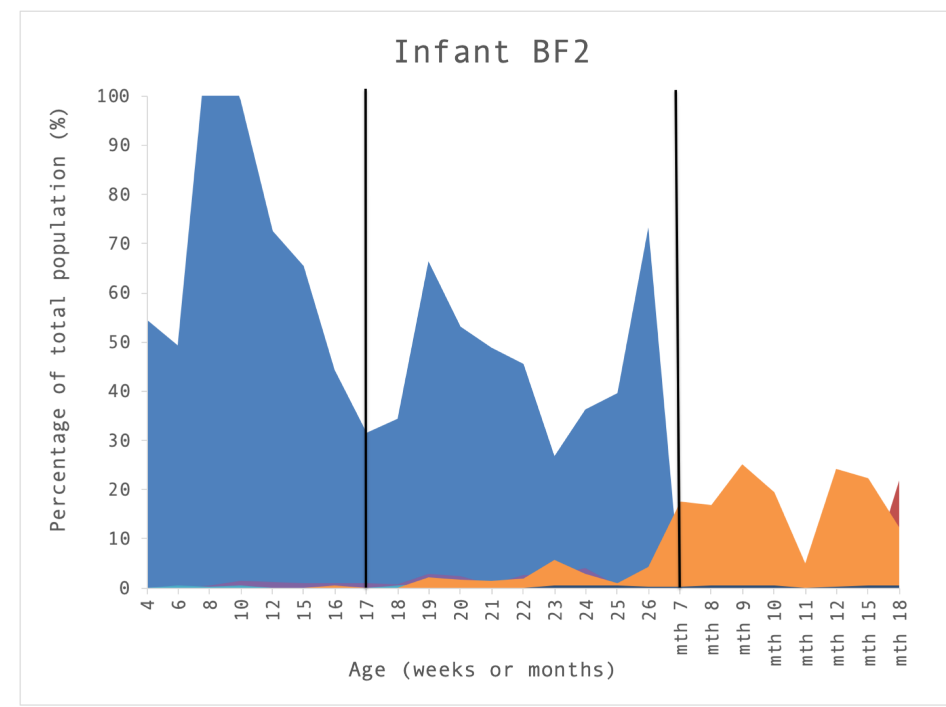
Journal Pre-proof

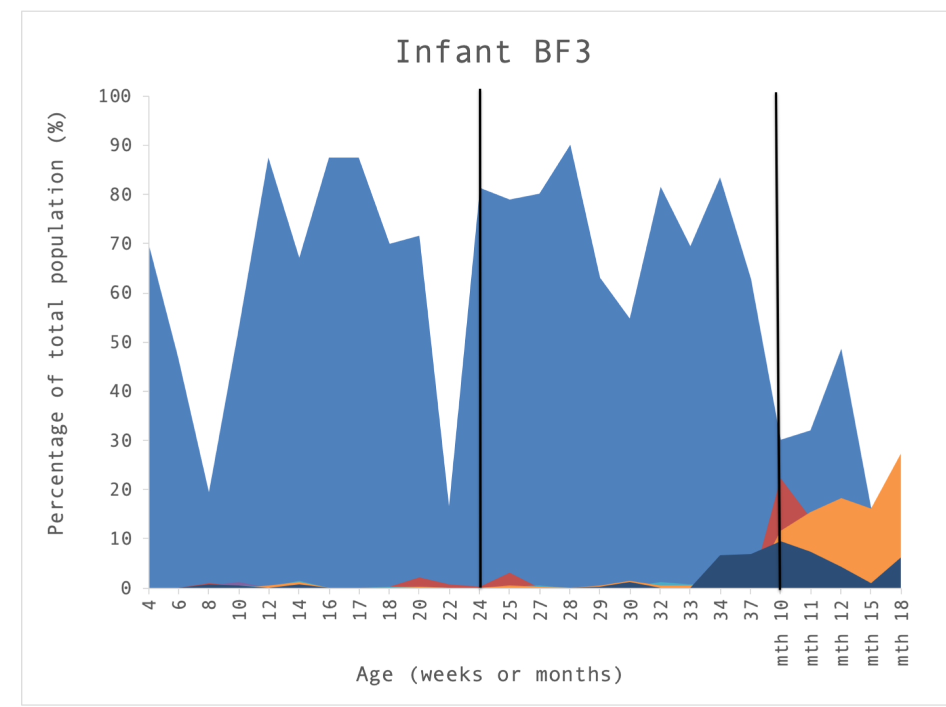
1087 supplemented with cysteine HCl at 0.05% and 2% (w/v) of selected carbohydrates, except
1088 for pectin and mucin (1%). Experiments were performed in triplicates; the data are
1089 presented as difference in OD between T_2 (30min) and T_{end} (48h) for growth on 12 selected
1090 carbohydrate sources and T_2 (30min) and T_{end} (72h) for an additional experiment on
1091 cellobiose (separate tabs). Related to Figure 5.
1092
1093 **Table S13.** Proteomics data from growth of strain B_25 on LNnT and cellobiose, and strain
1094 B_71 on 2'-FL (separate tabs). Related to Figure 6.
1095
1096 **Table S14.** Homology of proteins identified by proteomics in strain B_25 grown on
1097 cellobiose and LNnT, and strain B_71 grown on 2'-FL to known *Bifidobacterium* gene clusters
1098 (BLAST+, e-value 1e-50) (separate tabs). Related to Figure 6.
1099
1100

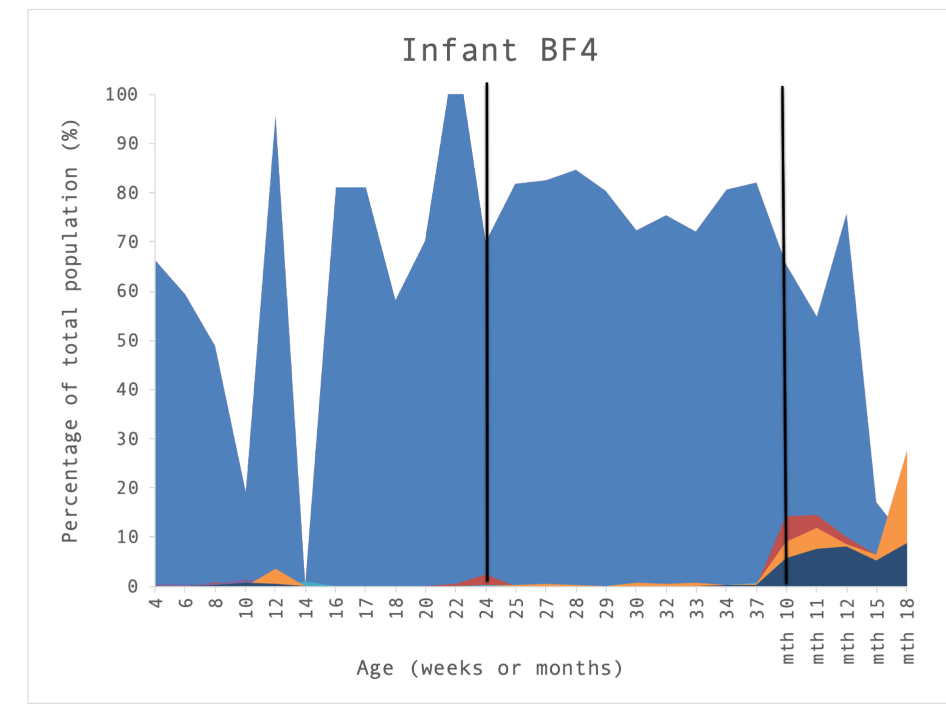
Journal Pre-proof

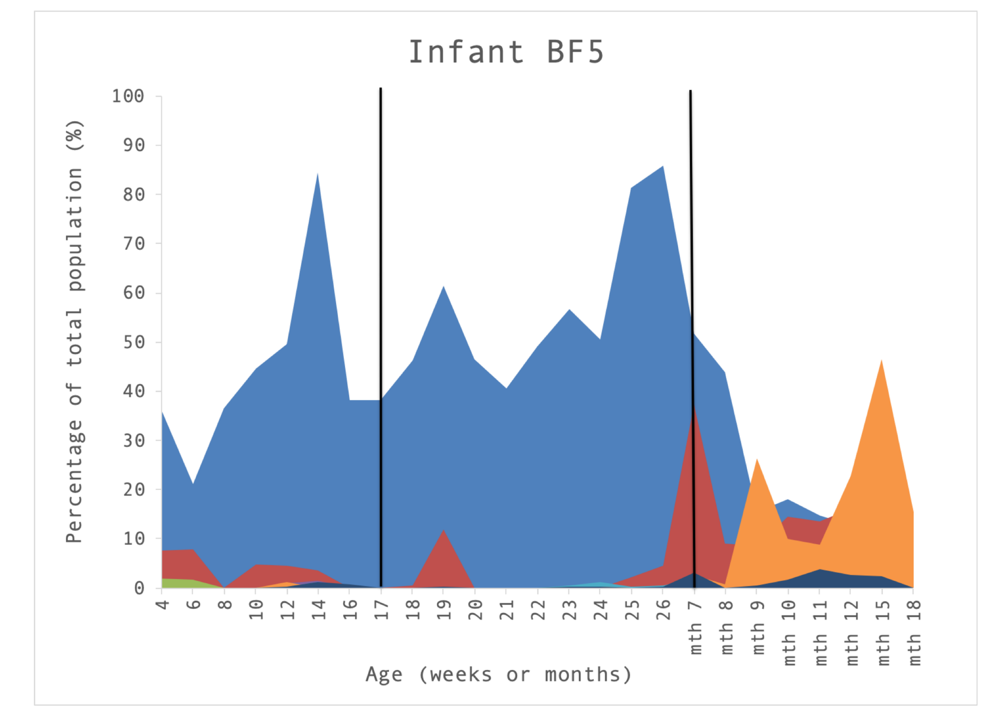
a)



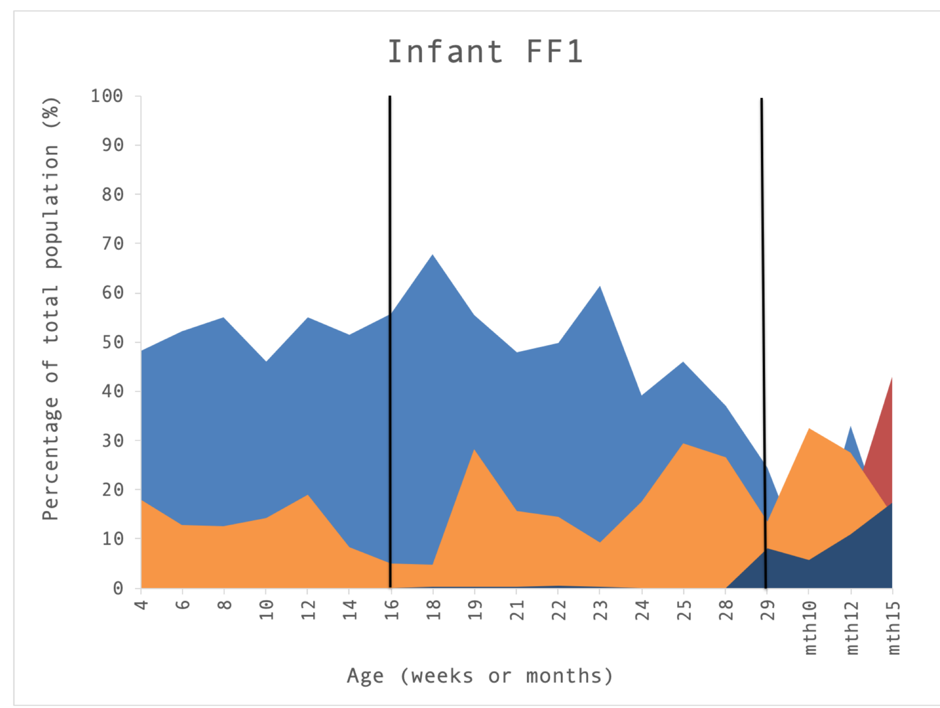


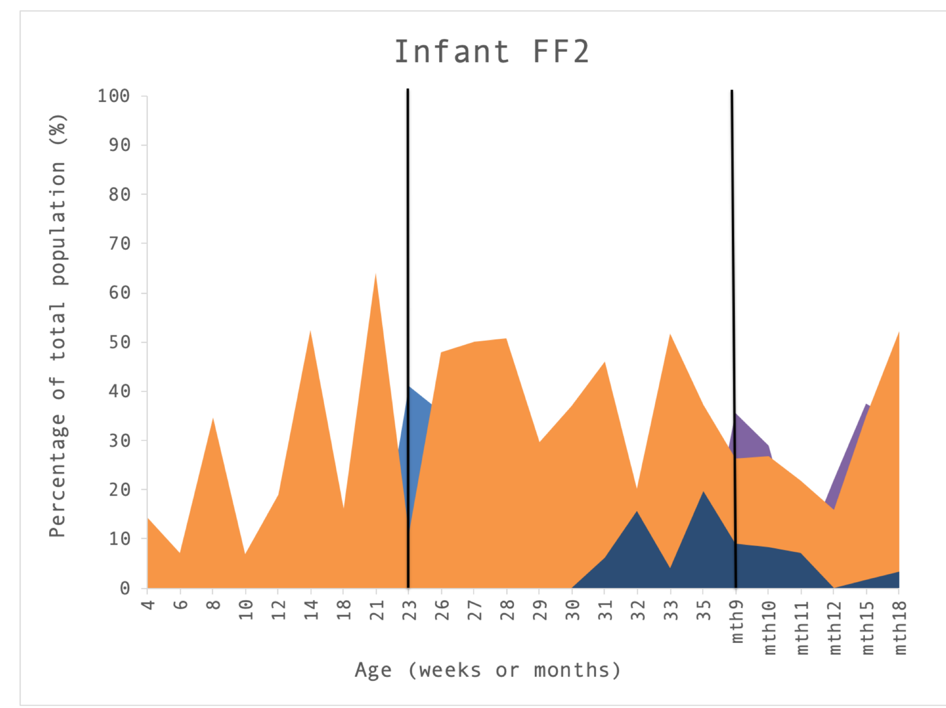


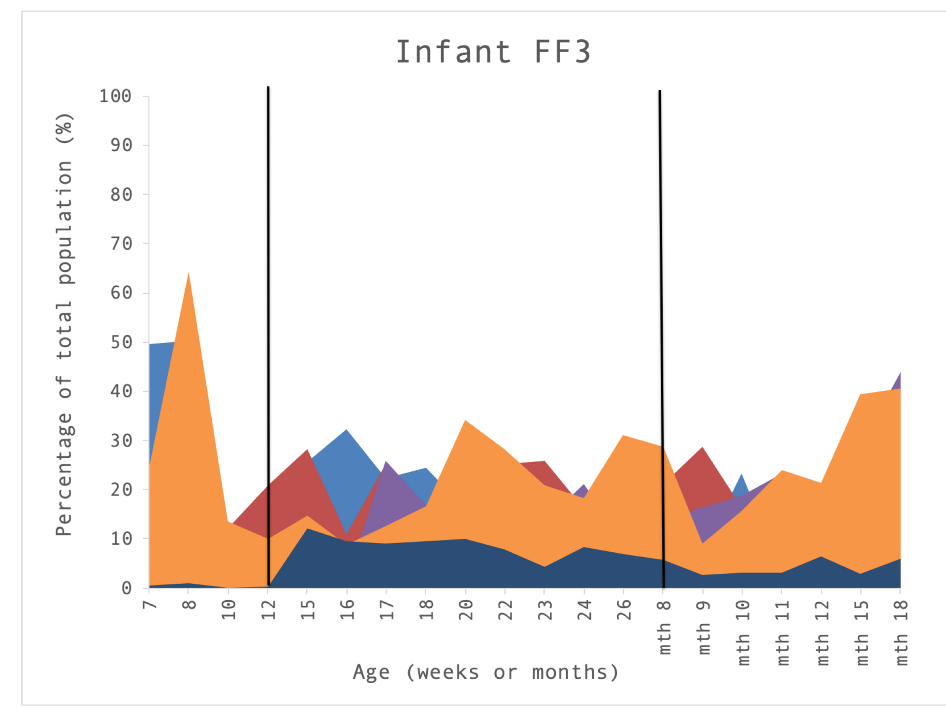


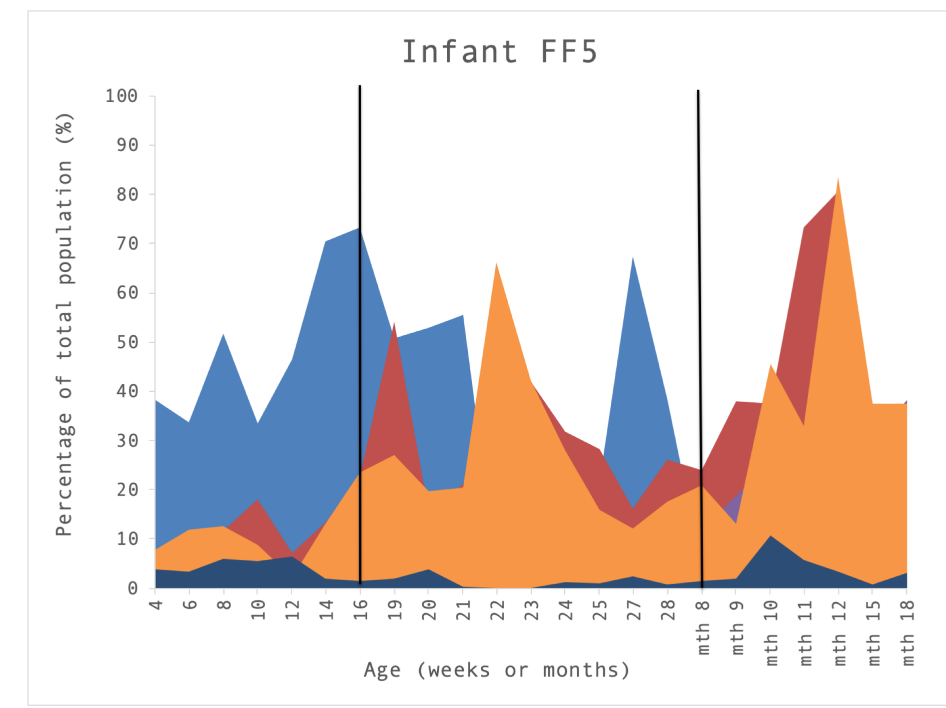


b)

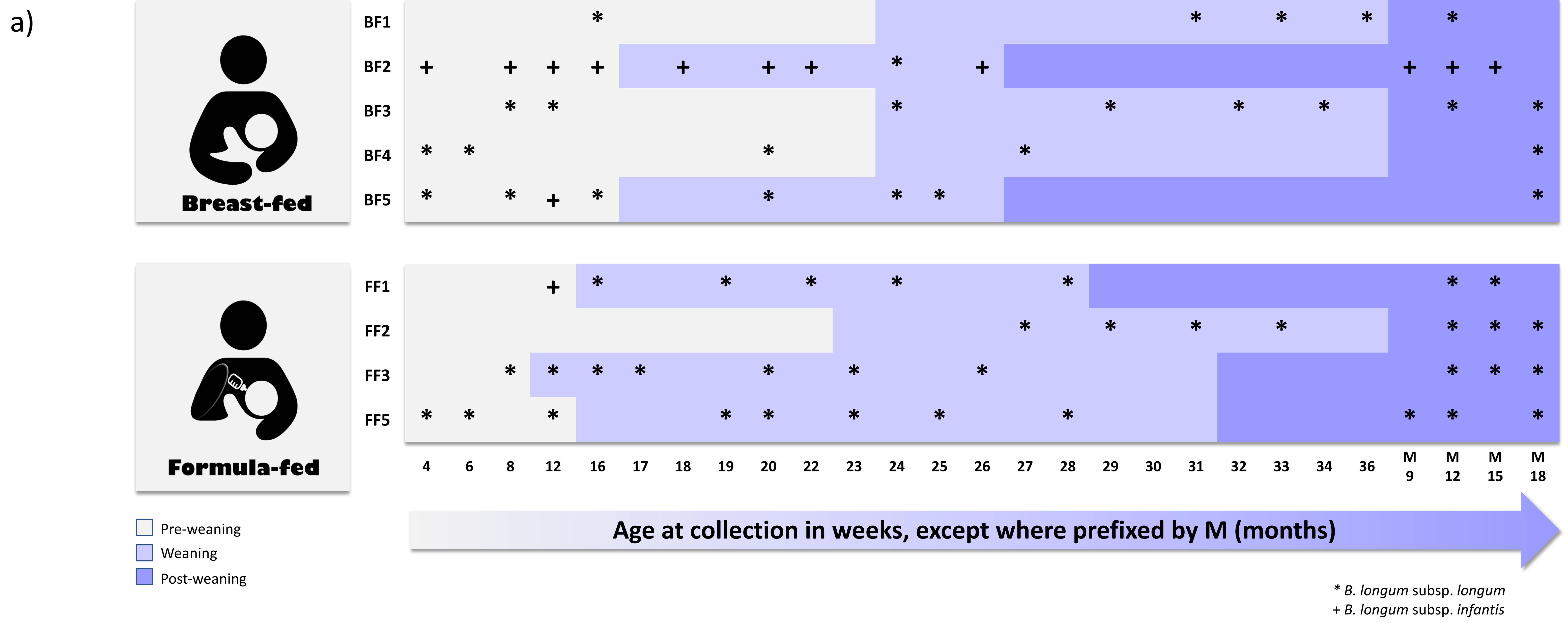








a)



* *B. longum* subsp. *longum*
+ *B. longum* subsp. *infantis*

b)

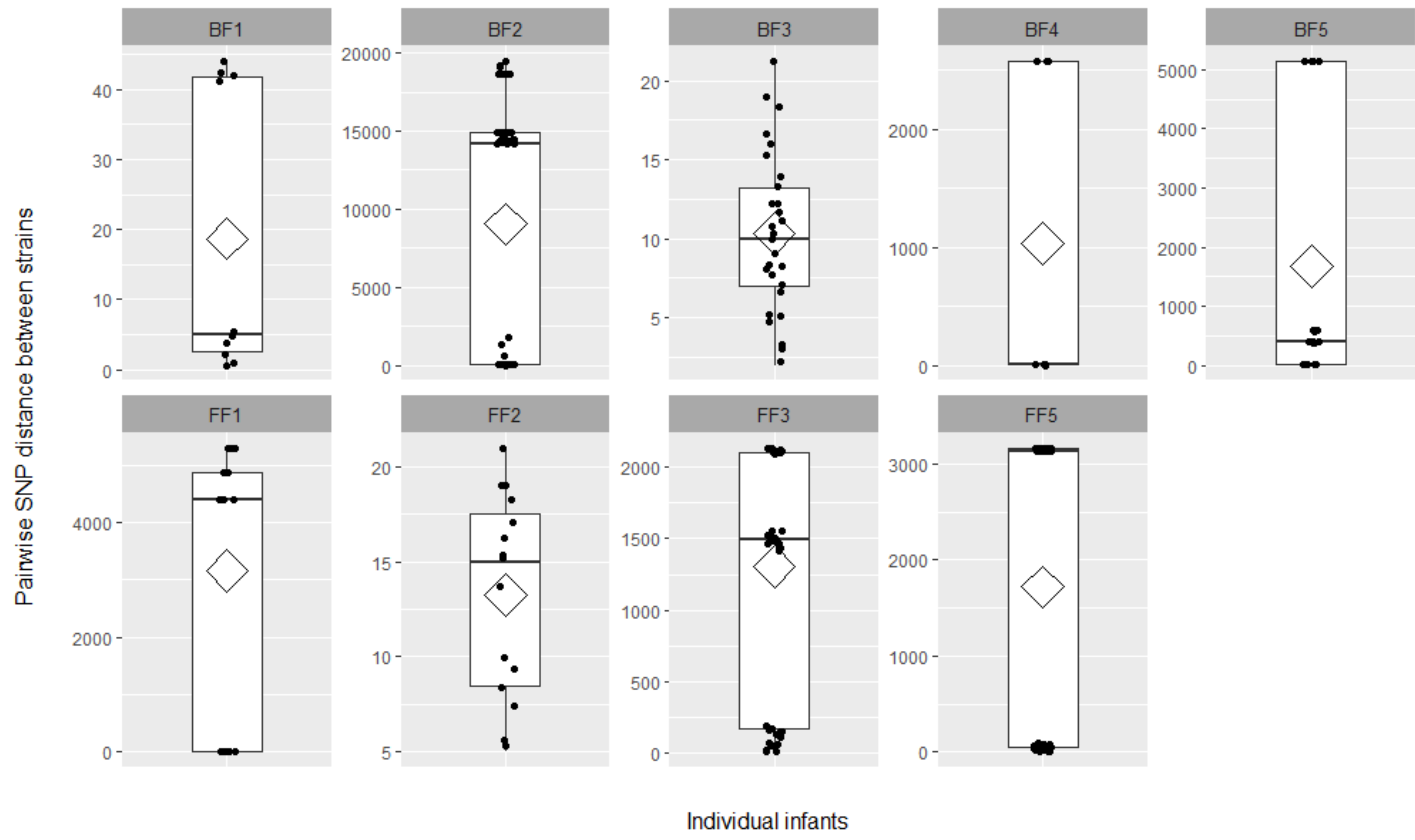
Tree scale: 0.001

Isolation period
Pre-weaning
Weaning
Post-weaning

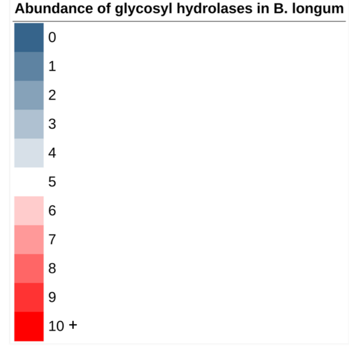
Isolation source
BF1
BF2
BF3
BF4
BF5
FF1
FF2
FF3
FF5

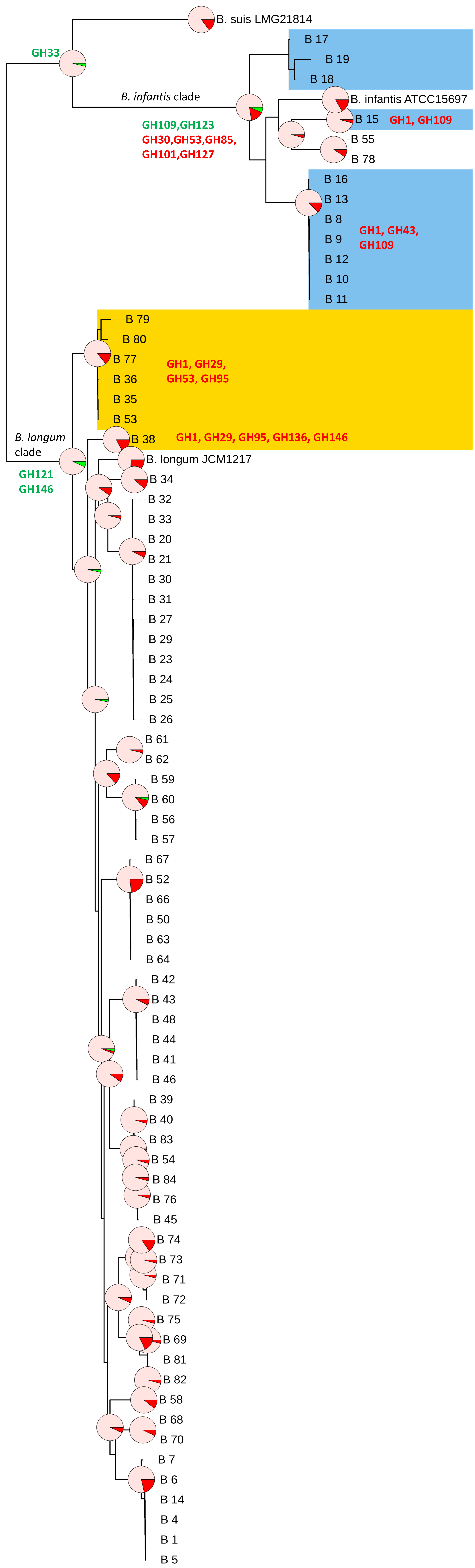
Isolation period
Isolation source

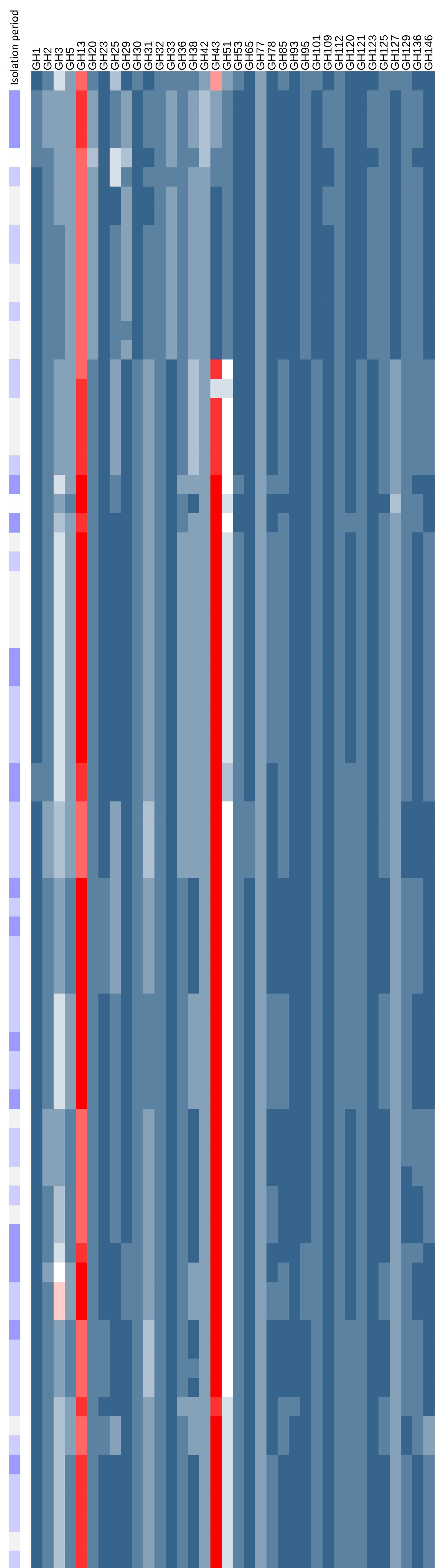
B longum subsp suis LMG21814
B 17
B 19
B 18
B longum subsp infantis ATCC15697
B 15
B 55
B 78
B 16
B 13
B 9
B 10
B 8
B 12
B 11
B 79
B 80
B 36
B 77
B 53
B 35
B 34
B 20
B 26
B 23
B 25
B 24
B 21
B 33
B 27
B 29
B 30
B 32
B 31
B longum subsp longum JCM1217
B 61
B 62
B 56
B 60
B 59
B 57
B 38
B 42
B 43
B 46
B 48
B 41
B 44
B 64
B 50
B 63
B 66
B 52
B 67
B 58
B 7
B 14
B 5
B 4
B 1
B 6
B 39
B 83
B 40
B 54
B 76
B 45
B 84
B 68
B 70
B 75
B 69
B 81
B 82
B 74
B 73
B 71
B 72

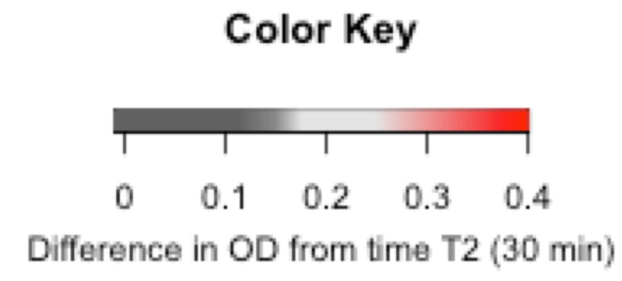

Summary of SNP differences between B. longum strains within individual babies
BF1
BF2
BF3
BF4
BF5
FF1
FF2
FF3
FF5
Pairwise SNP distance between strains
Individual infants

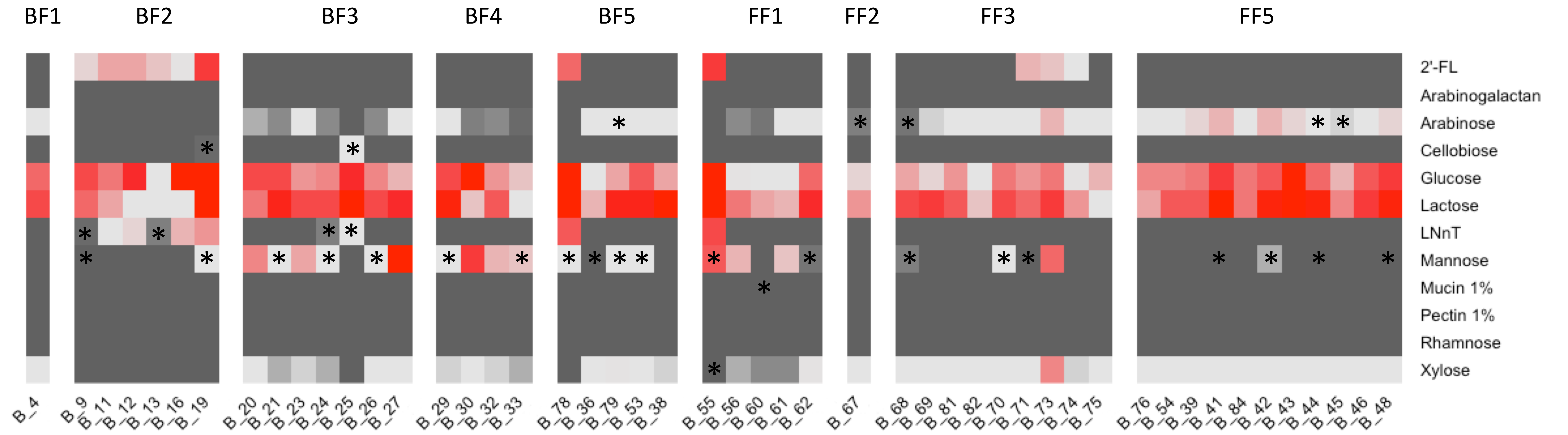
Tree scale: 0.01

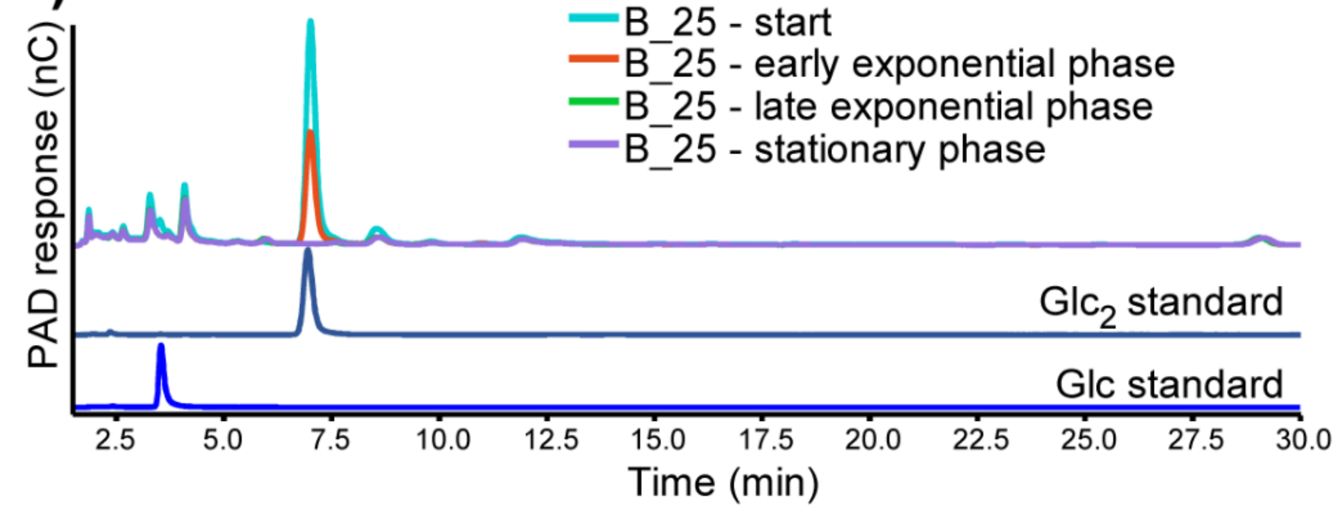


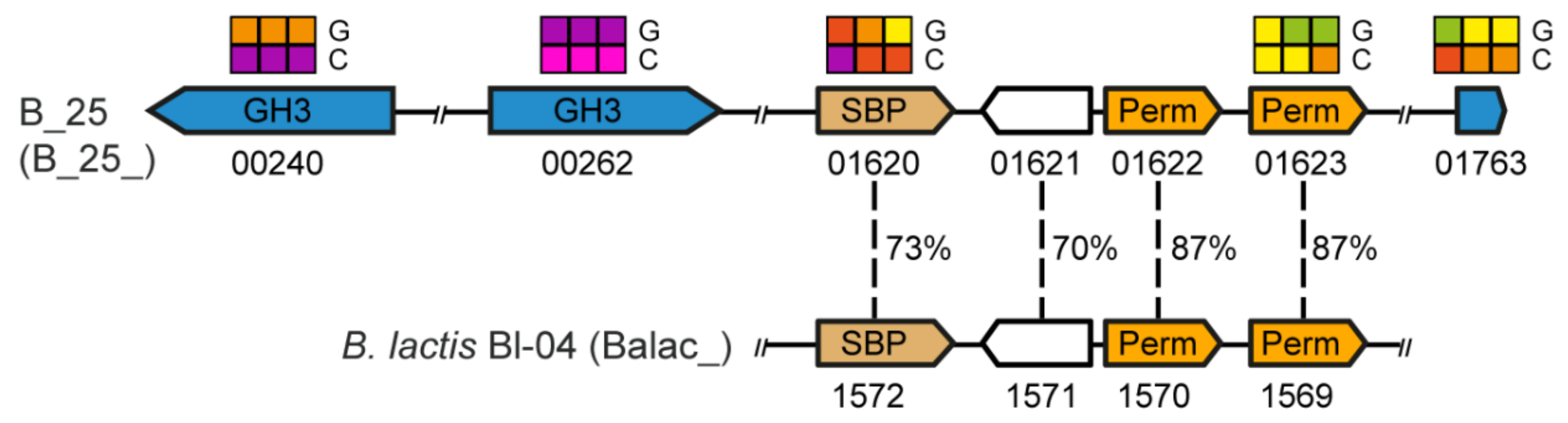


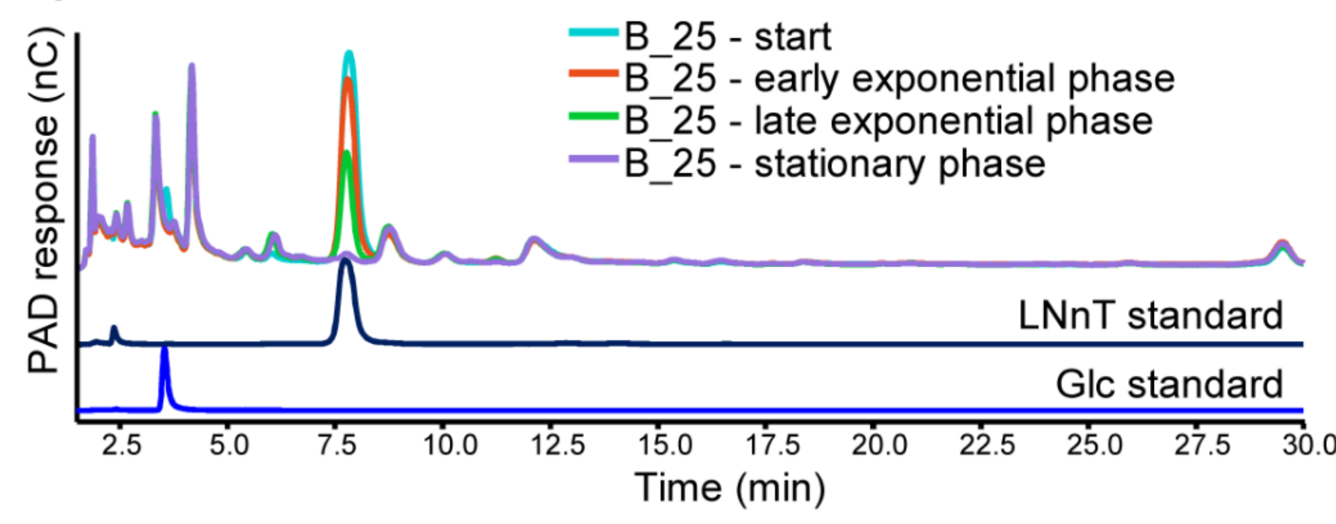


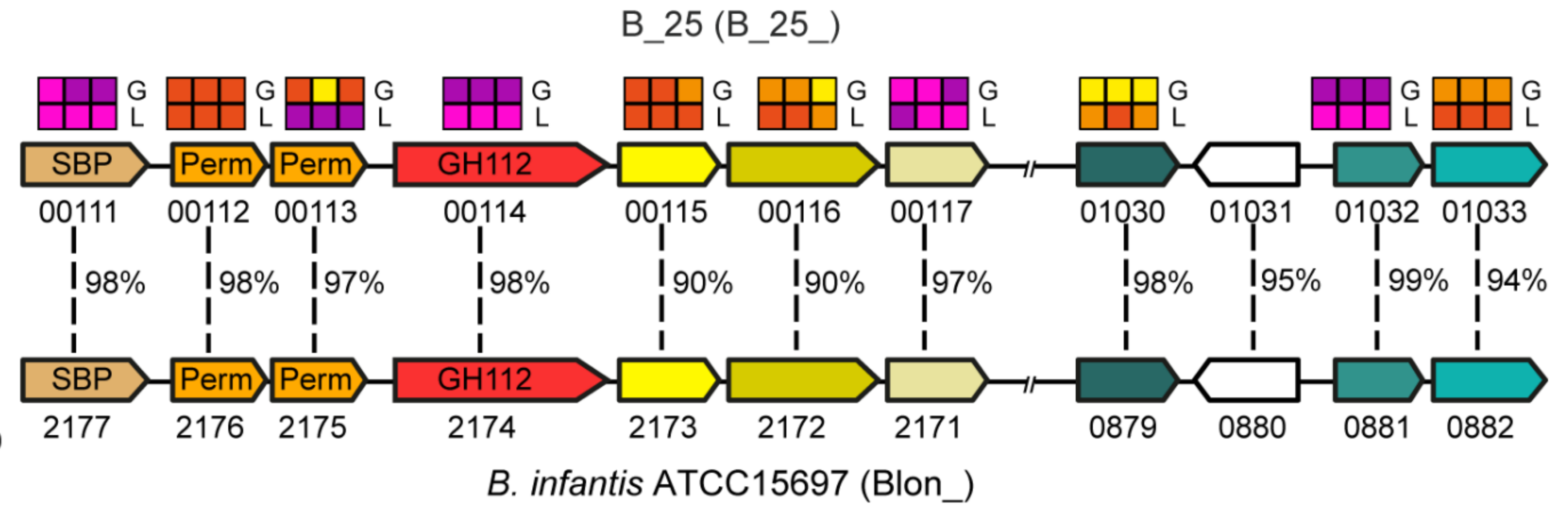

Color Key
0
0.1
0.2
0.3
0.4
Difference in OD from time T2 (30 min)

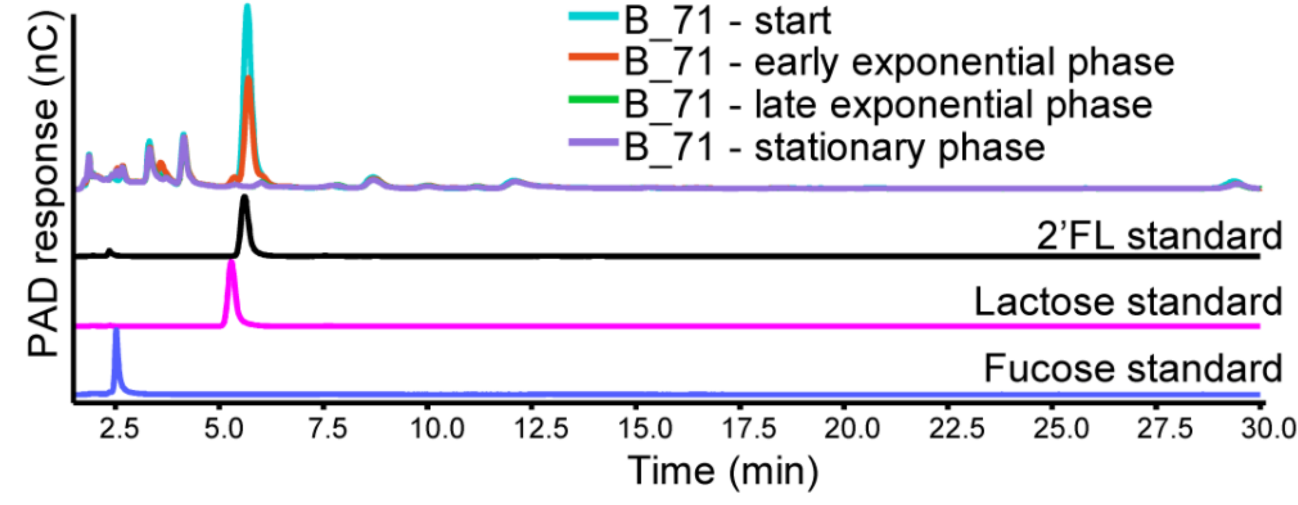

BF1
BF2
BF3
BF4
BF5
FF1
FF2
FF3
FF5
2'-FL
Arabinogalactan
Arabinose
Cellobiose
Glucose
Lactose
LNnT
Mannose
Mucin 1%
Pectin 1%
Rhamnose
Xylose
B_4
B_9
B_11
B_12
B_13
B_16
B_19
B_20
B_21
B_23
B_24
B_25
B_26
B_27
B_29
B_30
B_32
B_33
B_78
B_36
B_79
B_53
B_38
B_55
B_56
B_60
B_61
B_62
B_67
B_68
B_69
B_81
B_82
B_70
B_71
B_73
B_74
B_75
B_76
B_54
B_39
B_41
B_84
B_42
B_43
B_44
B_45
B_46
B_48

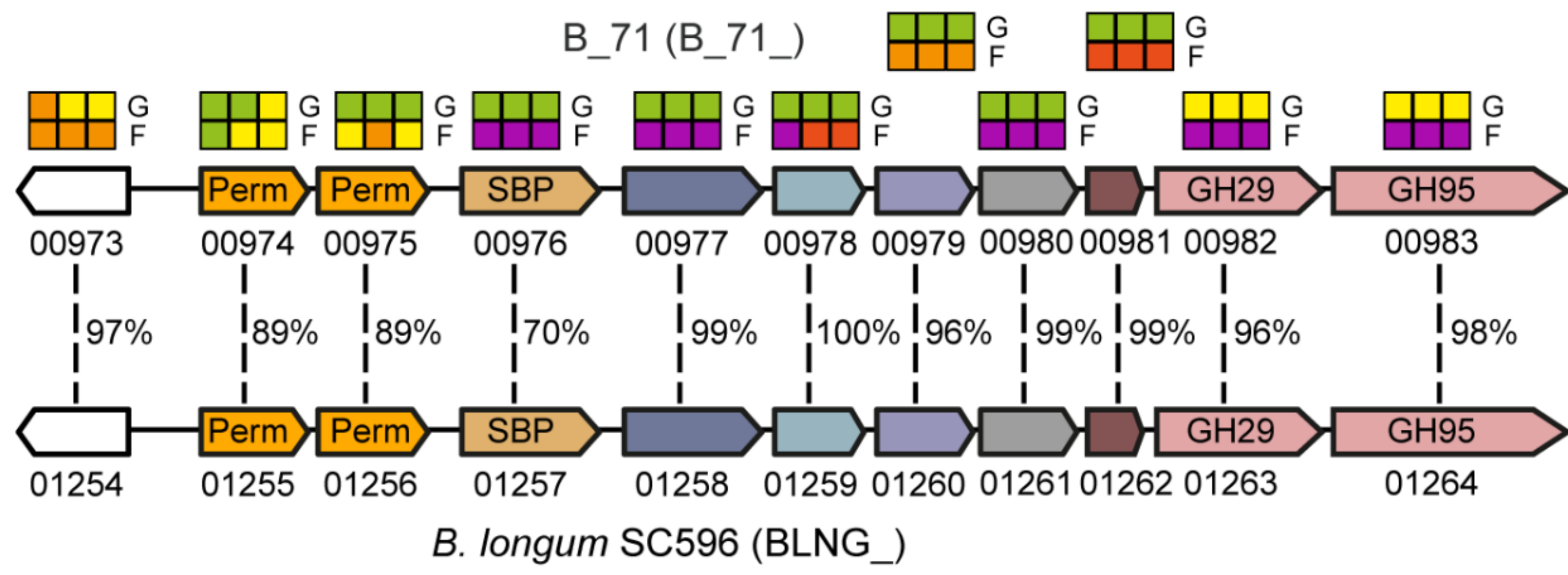

a) cellobiose
B_25 - start
B_25 - early exponential phase
B_25 - late exponential phase
B_25 - stationary phase
PAD response (nC)
Glc2 standard
Glc standard
Time (min)


B_25
(B_25_)
GH3
00240
GH3
00262
SBP
01620
01621
Perm
01622
Perm
01623
01763
73%
70%
87%
87%
B. lactis Bl-04 (Balac_)
SBP
1572
1571
Perm
1570
Perm
1569

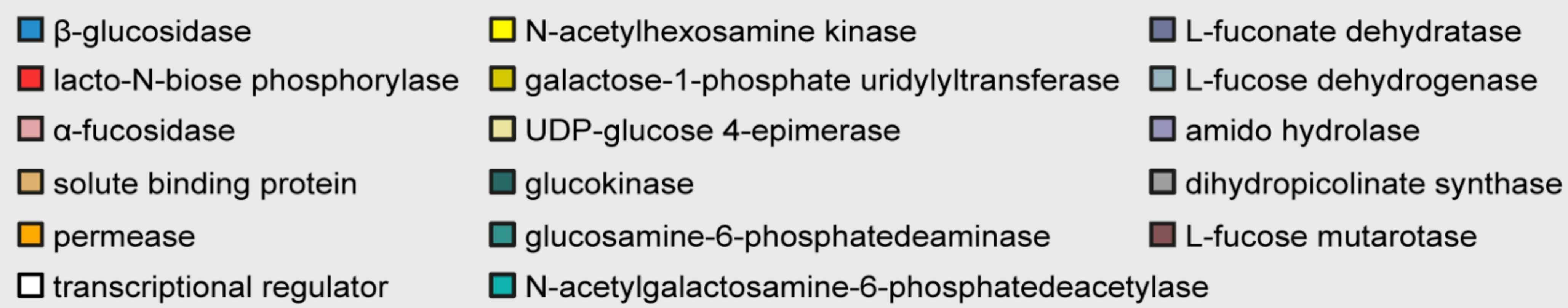

b) LNnT
B_25 - start
B_25 - early exponential phase
B_25 - late exponential phase
B_25 - stationary phase
PAD response (nC)
LNnT standard
Glc standard
Time (min)


B_25 (B_25_)
SBP
00111
Perm
00112
Perm
00113
GH112
00114
00115
00116
00117
01030
01031
01032
01033
98%
98%
97%
98%
90%
90%
97%
98%
95%
99%
94%
2177
2176
2175
2174
2173
2172
2171
0879
0880
0881
0882
B. infantis ATCC15697 (Blon_)


c) 2'FL
B_71 - start
B_71 - early exponential phase
B_71 - late exponential phase
B_71 - stationary phase
PAD response (nC)
2'FL standard
Lactose standard
Fucose standard
Time (min)


B_71 (B_71_)
00973
Perm
00974
Perm
00975
SBP
00976
00977
00978
00979
00980
00981
GH29
00982
GH95
00983
97%
89%
89%
70%
99%
100%
96%
99%
99%
96%
98%
01254
01255
01256
01257
01258
01259
01260
01261
01262
01263
01264
B. longum SC596 (BLNG_)

Log2(LFQ)
33-34
31-32
29-30
27-28
25-26
19-24


Predicted gene function
β-glucosidase
lacto-N-biose phosphorylase
α-fucosidase
solute binding protein
permease
transcriptional regulator
N-acetylhexosamine kinase
galactose-1-phosphate uridylyltransferase
UDP-glucose 4-epimerase
glucokinase
glucosamine-6-phosphatedeaminase
N-acetylgalactosamine-6-phosphatedeacetylase
L-fuconate dehydratase
L-fucose dehydrogenase
amido hydrolase
dihydropicolinate synthase
L-fucose mutarotase

Journal Pre-proof

Highlights:

- *B. longum* strains can persist in individual infants despite dietary changes
- New strains with additional metabolism capabilities can also be acquired
- *B. longum* carbohydrate metabolism repertoires are individual and strain-specific
- Strains at different dietary stages show genomic adaptations to specific substrates

Journal Pre-proof