Differentiation of *Bifidobacterium longum* subspecies *longum* and *infantis* by quantitative PCR using functional gene targets

Blair Lawley [Corresp., 1] , **Karen Munro** [1] , **Alan Hughes** [1] , **Alison J Hodgkinson** [2] , **Colin G Prosser** [3] , **Dianne Lowry** [3] , **Shao J Zhou** [4, 5] , **Maria Makrides** [4] , **Robert A Gibson** [5] , **Christophe Lay** [6] , **Charmaine Chew** [6] , **Pheng Soon Lee** [7] , **Khai Hong Wong** [7] , **Gerald W Tannock** [1, 8, 9]

[1] Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand

[2] Ruakura, AgResearch, Hamilton, New Zealand

[3] Dairy Goat Cooperative (NZ) Ltd., Hamilton, New Zealand

[4] Women's and Children's Health Research Institute, Adelaide, Australia

[5] School of Agriculture, Food and Wine, University of Adelaide, Adelaide, Australia

[6] Danone Nutricia Research, Biopolis, Singapore

[7] Mead Johnson Nutrition, Marina Bay, Singapore

[8] Riddet Institute Centre of Research Excellence, Palmerston North, New Zealand

[9] Microbiome Otago, University of Otago, Dunedin, New Zealand

Corresponding Author: Blair Lawley
Email address: blair.lawley@otago.ac.nz

Background. Members of the genus *Bifidobacterium* are abundant in the feces of babies during the exclusively-milk-diet period of life. *Bifidobacterium longum* is reported to be a common member of the infant fecal microbiota. However, *B. longum* is composed of three subspecies, two of which are represented in the bowel microbiota (*B. longum* subsp. *longum*; *B. longum* subsp. *infantis*). *B. longum* subspecies are not differentiated in many studies, so that their prevalence and relative abundances are not accurately known. This may largely be due to difficulty in assigning subspecies identity using DNA sequences of *16S rRNA* or *tuf* genes that are commonly used in bacterial taxonomy.

Methods. We developed a qPCR method targeting the sialidase gene (subsp. *infantis*) and sugar kinase gene (subsp. *longum*) to differentiate the subspecies using specific primers and probes. Specificity of the primers/probes was tested by *in silico,* pangenomic search, and using DNA from standard cultures of bifidobacterial species. The utility of the method was further examined using DNA from feces that had been collected from infants inhabiting various geographical regions.

Results. A pangenomic search of the NCBI genomic database showed that the PCR primers/probes targeted only the respective genes of the two subspecies. The primers/probes showed total specificity when tested against DNA extracted from the gold standard strains (type cultures) of bifidobacterial species detected in infant feces. Use of the qPCR method with DNA extracted from the feces of infants of different ages, delivery method and nutrition, showed that subsp. *infantis* was detectable (0-32.4% prevalence) in the feces of Australian (n = 90), South-East Asian (n = 24), and Chinese babies (n = 91), but in all cases at low abundance (<0.01-4.6%) compared to subsp. *longum* (0.1-33.7% abundance; 21.4-100% prevalence).

Discussion. Our qPCR method differentiates *B. longum* subspecies *longum* and *infantis* using characteristic functional genes. It can be used as an identification aid for isolates of bifidobacteria, as well as in determining prevalence and abundance of the subspecies in feces. The method should thus be useful in ecological studies of the infant gut microbiota during early life where an understanding of the

ecology of bifidobacterial species may be important in developing interventions to promote infant health.

1   **Differentiation of *Bifidobacterium longum* subspecies *longum* and *infantis* by quantitative**

2   **PCR using functional gene targets.**

4   **Blair Lawley [1], Karen Munro [1], Alan Hughes [1], Alison J. Hodgkinson [2], Colin G. Prosser [3],**

5   **Dianne Lowry [3], Shao J. Zhou [4, 5], Maria Makrides [4], Robert A. Gibson [5], Christophe Lay [6],**

6   **Charmaine Chew [6], Pheng Soon Lee [7], Khai Hong Wong [7], Gerald W. Tannock [1, 8, 9].**

7   [1] Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand

8   [2] AgResearch Ruakura, Hamilton, New Zealand

9   [3] Dairy Goat Cooperative (NZ) Ltd., Hamilton, New Zealand

10   [4] Women's and Children's Health Research Institute, Adelaide, Australia

11   [5] School of Agriculture, Food, and Wine, University of Adelaide, Adelaide, Australia

12   [6] Danone Nutricia Research, Biopolis, Singapore

13   [7] Mead Johnson Nutrition, Marina Bay, Singapore

14   [8] Riddet Institute Centre of Research Excellence, Palmerston North, New Zealand

15   [9] Microbiome Otago, University of Otago, Dunedin, New Zealand

17   Corresponding author:

18   Blair Lawley [1]

19   Email address: blair.lawley@otago.ac.nz

21   **Abstract**

22   **Background.** Members of the genus *Bifidobacterium* are abundant in the feces of babies during

23   the exclusively-milk-diet period of life. *Bifidobacterium longum* is reported to be a common

24   member of the infant fecal microbiota. However, *B. longum* is composed of three subspecies,

25   two of which are represented in the bowel microbiota (*B. longum* subsp. *longum*; *B. longum*

26   subsp. *infantis*). *B. longum* subspecies are not differentiated in many studies, so that their

27   prevalence and relative abundances are not accurately known. This may largely be due to

28   difficulty in assigning subspecies identity using DNA sequences of *16S rRNA* or *tuf* genes that

29   are commonly used in bacterial taxonomy.

30   **Methods.** We developed a qPCR method targeting the sialidase gene (subsp. *infantis*) and sugar

31   kinase gene (subsp. *longum*) to differentiate the subspecies using specific primers and probes.

32   Specificity of the primers/probes was tested by *in silico,* pangenomic search, and using DNA

33   from standard cultures of bifidobacterial species. The utility of the method was further examined

34   using DNA from feces that had been collected from infants inhabiting various geographical

35   regions.

36   **Results.** A pangenomic search of the NCBI genomic database showed that the PCR

37   primers/probes targeted only the respective genes of the two subspecies. The primers/probes

38   showed total specificity when tested against DNA extracted from the gold standard strains (type

39   cultures) of bifidobacterial species detected in infant feces. Use of the qPCR method with DNA

40   extracted from the feces of infants of different ages, delivery method and nutrition, showed that

41   subsp. *infantis* was detectable (0-32.4% prevalence) in the feces of Australian (n = 90), South-

42   East Asian (n = 24), and Chinese babies (n = 91), but in all cases at low abundance (<0.01-4.6%)

43   compared to subsp. *longum* (0.1-33.7% abundance; 21.4-100% prevalence).

44   **Discussion.** Our qPCR method differentiates *B. longum* subspecies *longum* and *infantis* using

45   characteristic functional genes. It can be used as an identification aid for isolates of

46   bifidobacteria, as well as in determining prevalence and abundance of the subspecies in feces.

47   The method should thus be useful in ecological studies of the infant gut microbiota during early

48   life where an understanding of the ecology of bifidobacterial species may be important in

49   developing interventions to promote infant health.

50

51 **INTRODUCTION**

52 Bifidobacteria commonly dominate the fecal microbiota of infants during the exclusively milk-

53 fed period of life (Biavati et al., 1984; Harmsen et al., 2000; Favier et al., 2002; Young et al.,

54 2004; Mariat et al., 2009; Roger et al., 2010; Coppa et al., 2011; Grönlund et al., 2011; Tannock

55 et al., 2012; Turroni et al., 2012; Yatsunenko et al., 2012; Makino et al., 2013; Huda et al., 2014;

56 Bäckhed et al., 2015; Milani et al., 2015; Vazquez-Gutierrez et al., 2015; Lewis et al., 2015;

57 Martin et al., 2016). This observation can be at least partly explained by the fact that

58 bifidobacterial species that are enriched in the infant bowel can utilize Human Milk

59 Oligosaccharides (HMO) or their components as growth substrates (Garrido, Dallas & Mills,

60 2013). *B. longum* is commonly detected in studies of the infant fecal microbiota. However, *B.*

61 *longum* is composed of three subspecies, two of which are represented in the bowel microbiota

62 (*B. longum* subsp. *longum*; *B. longum* subsp. *infantis*). Differentiation of *B. longum* into

63 subspecies has been carried out in relatively few quantitative studies of the infant fecal

64 microbiota, so that subspecies prevalence and relative abundances are not accurately known on a

65 global scale (Grönlund et al., 2011; Huda et al., 2014; Martin et al., 2016). This is probably due

66 to the difficulty of differentiating the *B. longum* subspecies using *16S rRNA* gene sequences

67 (Youn, Seo & Ji, 2008). Previously, we tested primers targeting the *16S rRNA* gene and found

68 weak cross-reactivity between subspecies when high concentrations of non-target template were

69 present (Tannock et al., 2012). Other groups have used *16S rRNA* gene primers (Matsuki et al.,

70 1999), 16S/23S intergenic region primers (Haarman & Knol, 2006), *23S rRNA* gene primers

71 (Kurakawa et al., 2015) and *tuf* gene primers (Sheu et al., 2010) to differentiate the subspecies,

72 but a high degree of similarity between sequences of marker genes, such as the *16S rRNA* gene

73 and *tuf* gene, leads to difficulty in generating PCR primers that are capable of absolute

74   discrimination of the *B. longum* subspecies (Youn, Seo & Ji, 2008). Inaccuracies in taxonomic

75   assignment (O'Callaghan et al., 2015) may also contribute to the difficulty in designing

76   discriminatory primers and/or probes based on all available sequence data. We believe an

77   approach targeting genes that define the functional differences between *B. longum* subspecies,

78   rather than phylogenetic marker genes, may be more appropriate and describe here a qPCR

79   method that distinguishes subsp. *longum* from subsp. *infantis* by targeting functional genes

80   exclusive to each of these subspecies. The method was used to determine the prevalence and

81   relative abundances of these two subspecies in the feces of children in Australia, South-East Asia,

82   and China that were breast milk or formula-fed, or delivered vaginally or by caesarean section.

83

84   **MATERIALS AND METHODS**

85   **Quantitative PCR differentiation of *B. longum* subsp. *longum* and subsp. *infantis.***

86   Primers and Taqman® probes were designed to target the *B. longum* subsp. *infantis* sialidase gene

87   (Blon2348 from strain ATCC 15697) and the *B. longum* subsp. *longum* sugar kinase gene

88   (BL0274 from strain NCC 2705) using 'primer 3' (Untergasser et al., 2012). Predicted primers

89   were checked for target specificity in-silico using Primer-BLAST

90   (https://www.ncbi.nlm.nih.gov/tools/primer-blast/), and showed absence of predicted cross

91   reactivity with non-target templates. Primers and probes were obtained from IDT (Singapore)

92   and are described in Table 1. The primer/probe combinations were tested for reaction efficiency

93   and specificity using genomic DNA (gDNA) purified from bifidobacterial type cultures (the gold

94   standard cultures for species) of species reported to be detected in infant feces: *B. adolescentis*

95   (DSM 20083[T]), *B. animalis* subsp. *lactis* (DSM 10140[T]), *B. angulatum* DSM 20098[T], *B. bifidum*

96   (DSM 20456[T]), *B. breve* (ATCC 15700[T]), *B. catenulatum* (DSM 20224[T]), *B. dentium* (ATCC

97   27534^T), *B. longum* subsp. *infantis* (DSM 20088^T), *B. longum* subsp. *longum* (ATCC 15707^T), *B.*

98   *pseudocatenulatum* (DSM 20438^T), and *B. pseudolongum* (ATCC 25526^T), (Grönlund et al.,

99   2011; Makino et al., 2013; Huda et al., 2014; Bäckhed et al., 2015; Vazquez-Gutierrez et al.,

100   2015; Martin et al., 2016). A Life Technologies ViiA™7 real time PCR system and MicroAmp

101   Fast optical 96-well or 384-well plates with optical adhesive film (Applied Biosystems, Carlsbad,

102   CA) were used. All reactions were carried out in a final volume of 15 μl containing 1× TaqMan®

103   Fast PCR mastermix (Applied Biosystems), 300 nM of each primer and 100 nM TaqMan® probe.

104   For specificity testing, template DNA was diluted to 5 ng/μl, and 2 ng was added to each reaction.

105   The thermocycling profile consisted of an initial activation of the polymerase at 95°C for 30 s,

106   followed by 40 cycles of 95°C for 10 s and 60°C for 30 s. Fluorescence levels were measured

107   after the 60°C annealing/extension step. Standard curves (to measure reaction efficiency) were

108   generated using gDNA extracted from bifidobacterial strains *Bifidobacterium longum* subsp.

109   *longum* (ATCC 15707^T) and *Bifidobacterium longum* subsp. *infantis* (DSM 20088^T) using the

110   bead-beating phenol/chloroform/ethanol protocol described previously (Tannock et al., 2012).

111   The standard DNA was quantified spectrophotometrically using a NanoDrop 1000

112   spectrophotometer (Thermo Scientific) and diluted in 10-fold steps from 5×10^6 to 5×10^1

113   genomes/reaction, calculated using target gene copies per genome obtained from genome

114   sequence information (NCBI). All reactions were carried out in duplicate and were run twice on

115   separate plates. No-template controls were also included on each plate.

116   **Determination of the prevalence and abundances of subspecies *longum* and *infantis* in**

117   **infant feces.**

118   Standard curve-based qPCR was used to determine the relative abundance and prevalence of *B.*

119   *longum* subspecies. The CFU/ml of *B. longum* subsp. *longum* (ATCC 15707^T) and *B. longum*

120   subsp. *infantis* (DSM 20088^T) was determined by spreading dilutions of 7-hour anaerobic

121   cultures (Lactobacilli MRS broth, Difco) on Lactobacilli MRS agar to obtain colony counts after

122   48 hours incubation. Genomic DNA was extracted from 1 ml of culture using the bead-beating

123   phenol/chloroform/ethanol protocol described previously (Tannock et al., 2012). Standard curves

124   were generated with template quantities equivalent to 5 x 10^6 to 5 x 10^1 CFU in a 10-fold

125   dilution series. Total community *16S rRNA* gene targets were determined using universal primers

126   and a SYBR® green detection system as previously described (Tannock et al., 2012) whilst *B.*

127   *longum* subspecies target numbers were obtained using the subspecies-specific Taqman® assays.

128   Relative abundance of each subspecies was determined by dividing the subspecies target quantity

129   by the total community target quantity and multiplying by 100. If a *B. longum* subspecies assay

130   was positive, the sample was assumed to be positive for this subspecies. Prevalence was

131   determined by dividing the number of fecal samples positive for each subspecies by the total

132   number of samples per infant group, and multiplying by 100.

133   **DNA extraction from infant feces.**

134   The methodology for DNA extraction from feces of Australian and Chinese children was

135   described by Tannock et al. (Tannock et al., 2012) and included bead-beating to disrupt bacterial

136   cells, phenol-chloroform treatment, and ethanol precipitation. South-East Asian fecal DNA was

137   extracted from feces using QIAGEN DNA Stool Mini-Kit (QIAGEN, Hilden, Germany)

138   modified with bead-beating. In brief, approximately 200 mg of 0.1 mm glass beads were added

139   to 200 mg of stool sample and suspended in QIAGEN ASL buffer. Bead-beating used a

140   FastPrep-24 (M.P. Biomedicals, USA) for 3 repetitions of 1 min bead-beating with 5 min.

141   incubation on ice. Samples were then heated at 95°C for 15 min before centrifugation at 20,000 x

142   *g* for 1 min. The supernatant was transferred to a clean tube containing an InhibitEX tablet and

143   vortexed to mix. The manufacturer's instructions were followed thereafter. Extracted fecal DNA

144   were eluted in 50 µl AE buffer, checked on NanoDrop 2000 and stored at -20°C prior to analysis.

145   **Effect of DNA extraction method on detection of *B. longum* subspecies.**

146   As two distinct gDNA extraction methods were used in this study, we determined whether the

147   choice of extraction methodology would significantly impact on the quantitative detection of *B.*

148   *longum* subspecies. We extracted gDNA from one Australian and three Chinese infant fecal

149   samples using the original bead-beating/phenol-chloroform/ethanol precipitation method and two

150   commercial silica membrane-based extraction methods. The two commercial methods were the

151   QIAGEN DNA Stool Mini-Kit (with modifications described above) and the MoBio PowerSoil®

152   DNA isolation kit (MoBio, Carlsbad CA) which was used according to the manufacturer's

153   instructions. Extracted gDNA was used as template in separate qPCR reactions including the

154   universal *16S rRNA* gene primers (to determine total community target quantity) and each of the

155   *B. longum* subspecies Taqman® assay primer/probe sets (to determine target quantity for each *B.*

156   *longum* subspecies). Single aliquots of each fecal sample were extracted by each method and

157   qPCR reactions were carried out in duplicate as described above.

158   **Australian infants.**

159   The Australian infants included in this study were part of a larger study (Australian New Zealand

160   Clinical Trials Registry ACTRN12608000047392) (Zhou et al., 2014) in South Australia,

161   comparing growth and nutritional status of infants fed either goat milk-based infant formula or

162   cow milk-based infant formula (Dairy Goat Co-operative [NZ] Ltd., Hamilton, New Zealand).

163   Healthy term infants were recruited to a multicenter, double-blind, controlled feeding trial.

164   Infants were randomly allocated (stratified by sex and study center) to receive either goat milk or

165   cow milk formula before they were 2 weeks of age. Infants were exclusively fed the study

166 formulas (with no other liquids or solids except for water, vitamin or mineral supplements, or

167 medicines). Groups of 30 infants were randomly selected from the main study for fecal

168 microbiota analysis, and a parallel group of exclusively breast milk-fed infants was included as a

169 reference group (30 infants). Fecal samples were obtained when the children were 2 months of

170 age. Ethical approval for the work contained in this article was obtained from the WCHN Human

171 Research Ethics Committee. Other details of the infants are given in Tannock et al. (Tannock et

172 al., 2012).

173 **South-East Asian infants.**

174 The infants (various ethnicities) were part of a larger clinical study registered under the Dutch

175 Trial Register (NTR 2838). The infants were recruited as part of a multi-country, exploratory,

176 randomized double-blind, controlled study in Singapore (n=19) and Thailand (n=5). An equal

177 distribution of vaginally and elective caesarean section delivered infants were selected from each

178 location. The infants were 'mixed fed' (breast-milk and/or cow milk-based infant formula) from

179 birth. Fecal samples collected at Day 3/5 (hereafter termed < 1 week), 2 months and 4 months of

180 age were selected for DNA extraction and qPCR analysis.

181 **Chinese infants.**

182 The Chinese infants included in this study were part of Mead Johnson Nutrition Clinical Protocol

183 Number 8602. Healthy, singleton term infants, delivered by caesarian section, were recruited to

184 the trial. Infants (n = 40) were exclusively fed formula (with no other liquids or solids except for

185 water, vitamin or mineral supplements, or medicines). A parallel group of exclusively breast

186 milk-fed infants was included as a reference group (n = 51). Fecal samples were obtained when

187 the children were 6, 8, 10 and 12 weeks of age. Ethical approval for the work contained in this

188   article was obtained from the Ethics Committee, Xin Hua Hospital, affiliated to the School of

189   Medicine, Shanghai Jiao Tong University, China.

190

191   **RESULTS**

192   **Primer/probe specificity and efficiency.**

193   Reaction specificity and efficiency was tested by generating ten-fold serial dilutions of both *B.*

194   *longum* subspecies *longum* and *B. longum* subspecies *infantis* gDNA so that reactions would

195   contain between 5×10^6 to 5×10^1 genomes/reaction. For both of the primer/probe sets,

196   amplification product was only obtained for the appropriate specific target, and not in the non-

197   target template reactions (including those of other bifidobacterial species). Both primer/probe

198   sets achieved reaction efficiencies between 88% and 92% across four separate runs.

199   Modifications to annealing/extension temperatures, position of primers within the target gene

200   and primer concentration did not lead to improvements in reaction efficiency (data not shown).

201   Reaction efficiency was also tested in the presence of a complex mixture of fecal DNA lacking

202   bifidobacterial sequences. This DNA was obtained from the feces of an infant observed to be

203   free of bifidobacterial sequences as determined by *16S rRNA* gene (v1-3) high-throughput

204   sequencing. Addition of 1 ng/μl of this fecal genomic DNA did not impact on the efficiency of

205   either primer/probe set.

206   **Effect of differing genomic DNA extraction methods.**

207   Different DNA extraction methods were used in the infant study groups. Therefore, the potential

208   effect of extraction method was determined. Genomic DNA was extracted from the feces of four

209   infants using the phenol/chloroform/beadbeating method, the MoBio PowerSoil kit and the

210   Qiagen DNA Stool Mini kit. Results indicated that extraction methodology would have a

211 negligible impact on the detection of the *B. longum* subspecies (Table 2) but nevertheless argue

212 for a standard extraction method to be used in all studies.

213 **Prevalence and relative abundances of subspecies *longum* and *infantis* in feces.**

214 *B. longum* subsp. *infantis* and subsp. *longum* were detected in the feces of infants located in the

215 different geographical regions (Table 3). The prevalence of subspecies *infantis* tended to be less

216 than that of subspecies *longum*, and was mainly influenced by the age of the child rather than

217 nutrition; only two infants less than one week of age had detectable levels (<0.05% relative

218 abundance) of subspecies *infantis*.

219   In general, subspecies *infantis* was present in low abundance in infant feces compared to

220 *B. longum* subsp. *longum* (Table 4). The relative abundances of subspecies *infantis* were greatest

221 in Chinese samples and lowest in Australian children. In the case of South-East Asian and

222 Australian infants, caesarean section coupled with milk formula tended to be associated with low

223 abundances of subspecies *infantis* (Table 4). In cow's milk formula-fed Chinese infants, the

224 distribution of subspecies *longum* abundances tended to be different to those of breast milk-fed

225 infants (Figure 1). All of the Chinese infants had been delivered by caesarean, thus the

226 subspecies *longum* abundances were affected by infant nutrition. Overall, it was concluded that

227 subspecies *infantis* was prevalent, but was of low relative abundance compared to subspecies

228 *longum* in infant feces collected in all three geographical regions.

229

230 **DISCUSSION**

231 High-throughput sequencing of *16S rRNA* genes is a powerful approach when studying microbial

232 community composition (Caporaso et al., 2010), but is limited in its ability to accurately

233 discriminate very closely related bacteria. Sequences are clustered according to similarity but

234   this is influenced by multiple factors including the length of the region sequenced, its

235   evolutionary stability, and the error rate inherent to the sequencing chemistry used. Longer reads

236   allow finer taxonomic discrimination, but higher error rates lead to larger numbers of spurious

237   clusters, or operational taxonomic units (OTUs). The widely used clustering threshold of 97%

238   identity (an approximation to species level clusters) can lead to closely related organisms being

239   included in the same OTUs, as is the case with *B. longum* subspecies (Bäckhed et al., 2015). As

240   we demonstrate here, quantitative discrimination of bifidobacterial subspecies can be achieved

241   using specific primers and qPCR probes that target unique functional gene sequences. Our data

242   obtained from infants of different ethnicities and culture support the view that, of the two

243   subspecies, *longum* predominates in infant feces (Grönlund et al., 2011; Huda et al., 2014;

244   Martin et al., 2016). This might be unexpected, especially in the case of breast milk-fed infants

245   where the specialized HMO metabolism of subspecies *infantis* could favor proliferation of this

246   subspecies over subspecies *longum* (Sela et al., 2008). Investigations of the infant bowel

247   microbiota in terms of comparative bacterial growth using available substrates and resource

248   partitioning in the community are required to provide ecological explanations of bifidobacterial

249   relative abundances.

250   The evolutionary importance of the infant-bifidobacteria paradigm remains speculative,

251   but may involve competitive exclusion of pathogens and/or support of infant nutrition (Kunz et

252   al., 2000; D'Aimmo et al., 2012; Gordon et al., 2012; Ruhaak et al., 2014). The resolution of this

253   quandary requires an understanding of the roles of the different bifidobacterial species and

254   subspecies in the infant bowel, how they form co-operative consortia, and how their activities

255   impact on infant health. Knowledge of the relative proportions of the different kinds of

256   bifidobacteria is necessary for such studies to advance. Using our method, subspecies *longum*

257 and *infantis* can now be easily monitored in feces collected from infants born in different

258 locations and in different human societies where cultural practices may influence bowel ecology

259 in early life. This data may be important in developing interventions (such as the use of novel

260 oligosaccharides) to promote bifidobacterial consortia in the bowel that influence infant

261 development.

262

263 **CONCLUSION**

264 Bifidobacteria, represented by several species, are numerous in the feces of infants during the

265 exclusively milk-fed period of life. It is important to know the ecology of these species if health-

266 promoting interventions of the bowel microbiota are to proceed. We describe a method by which

267 two subspecies of *Bifidobacterium longum* can be differentiated quantitatively in infant

268 microbiota studies.

269

270 **Acknowledgements.**

271 We acknowledge Thomas Sullivan's contribution to the Australian study, and Jan Knol as

272 Danone global platform director Gut Biology and Microbiology.

273 **References.**

274

275 Bäckhed F., Roswall J., Peng Y., Feng Q., Jia H., Kovatcheva-Datchary P., Li Y., Xia Y., Xie H.,

276   Zhong H., Khan MT., Zhang J., Li J., Xiao L., Al-Aama J., Zhang D., Lee YS., Kotowska

277   D., Colding C., Tremaroli V., Yin Y., Bergman S., Xu X., Madsen L., Kristiansen K.,

278   Dahlgren J., Jun W. 2015. Dynamics and Stabilization of the Human Gut Microbiome

279   during the First Year of Life. *Cell Host and Microbe* 17:690–703. DOI:

280   10.1016/j.chom.2015.04.004.

281 Biavati B., Castagnoli P., Crociani F., Trovatelli LD. 1984. Species of the Bifidobacterium in the

282   feces of infants. *Microbiologica* 7:341–5.

283 Caporaso JG., Kuczynski J., Stombaugh J., Bittinger K., Bushman FD., Costello EK., Fierer N.,

284   Peña AG., Goodrich JK., Gordon JI., Huttley G a., Kelley ST., Knights D., Koenig JE., Ley

285   RE., Lozupone C a., McDonald D., Muegge BD., Pirrung M., Reeder J., Sevinsky JR.,

286   Turnbaugh PJ., Walters W a., Widmann J., Yatsunenko T., Zaneveld J., Knight R. 2010.

287   QIIME allows analysis of high-throughput community sequencing data. *Nature methods*

288   7:335–6. DOI: 10.1038/nmeth.f.303.

289 Coppa G V., Gabrielli O., Zampini L., Galeazzi T., Ficcadenti A., Padella L., Santoro L., Soldi

290   S., Carlucci A., Bertino E., Morelli L. 2011. Oligosaccharides in 4 different milk groups,

291   Bifidobacteria, and Ruminococcus obeum. *J Pediatr Gastroenterol Nutr* 53:80–87. DOI:

292   10.1097/MPG.0b013e3182073103.

293 D'Aimmo MR., Mattarelli P., Biavati B., Carlsson NG., Andlid T. 2012. The potential of

294   bifidobacteria as a source of natural folate. *Journal of Applied Microbiology* 112:975–984.

295   DOI: 10.1111/j.1365-2672.2012.05261.x.

296   Favier CF., Vaughan EE., De Vos WM., Akkermans ADL. 2002. Molecular monitoring of

297   succession of bacterial communities in human neonates. *Applied and Environmental*

298   *Microbiology* 68:219–226. DOI: 10.1128/AEM.68.1.219-226.2002.

299   Garrido D., Dallas DC., Mills DA. 2013. Consumption of human milk glycoconjugates by

300   infant-associated bifidobacteria: Mechanisms and implications. *Microbiology (United*

301   *Kingdom)* 159:649–664. DOI: 10.1099/mic.0.064113-0.

302   Gordon JI., Dewey KG., Mills D a., Medzhitov RM. 2012. The human gut microbiota and

303   undernutrition. *Science translational medicine* 4:137ps12. DOI:

304   10.1126/scitranslmed.3004347.

305   Grönlund M-M., Grześkowiak Ł., Isolauri E., Salminen S. 2011. Influence of mother's intestinal

306   microbiota on gut colonization in the infant. *Gut Microbes* 2:227–233. DOI:

307   10.4161/gmic.2.4.16799.

308   Haarman M., Knol J. 2006. Quantitative Real-Time PCR Analysis of Fecal Lactobacillus Species

309   in Infants Receiving a Prebiotic Infant Formula. 72:2359–2365. DOI:

310   10.1128/AEM.72.4.2359.

311   Harmsen HJ., Wildeboer-Veloo AC., Raangs GC., Wagendorp AA., Klijn N., Bindels JG.,

312   Welling GW. 2000. Analysis of intestinal flora development in breast-fed and formula-fed

313   infants by using molecular identification and detection methods. *Journal of pediatric*

314   *gastroenterology and nutrition* 30:61–7. DOI: 10.1097/00005176-200001000-00019.

315   Huda MN., Lewis Z., Kalanetra KM., Rashid M., Ahmad SM., Raqib R., Qadri F., Underwood

316   M a., Mills D a., Stephensen CB. 2014. Stool Microbiota and Vaccine Responses of Infants.

317   *Pediatrics*:peds.2013-3937-. DOI: 10.1542/peds.2013-3937.

318   Kunz C., Rudloff S., Baier W., Klein N., Strobel S. 2000. Oligosaccharides in human milk:

319   structural, functional, and metabolic aspects. *Annual review of nutrition* 20:699–722. DOI:

320   10.1146/annurev.nutr.20.1.699.

321 Kurakawa T., Ogata K., Tsuji H., Kado Y., Takahashi T., Kida Y., Ito M., Okada N., Nomoto K.

322   2015. Establishment of a sensitive system for analysis of human vaginal microbiota on the

323   basis of rRNA-targeted reverse transcription-quantitative PCR. *Journal of Microbiological*

324   *Methods* 111:93–104. DOI: 10.1016/j.mimet.2015.01.021.

325 Lewis ZT., Totten SM., Smilowitz JT., Popovic M., Parker E., Lemay DG., Van Tassell ML.,

326   Miller MJ., Jin Y-S., German JB., Lebrilla CB., Mills D a. 2015. Maternal

327   fucosyltransferase 2 status affects the gut bifidobacterial communities of breastfed infants.

328   *Microbiome* 3:1–21. DOI: 10.1186/s40168-015-0071-z.

329 Makino H., Kushiro A., Ishikawa E., Kubota H., Gawad A., Sakai T., Oishi K., Martin R., Ben-

330   Amor K., Knol J., Tanaka R. 2013. Mother-to-infant transmission of intestinal

331   bifidobacterial strains has an impact on the early development of vaginally delivered

332   infant's microbiota. *PLoS ONE* 8:e78331. DOI: 10.1371/journal.pone.0078331.

333 Mariat D., Firmesse O., Levenez F., Guimarães V., Sokol H., Doré J., Corthier G., Furet J-P.

334   2009. The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. *BMC*

335   *microbiology* 9:123. DOI: 10.1186/1471-2180-9-123.

336 Martin R., Makino H., Cetinyurek Yavuz A., Ben-Amor K., Roelofs M., Ishikawa E., Kubota H.,

337   Swinkels S., Sakai T., Oishi K., Kushiro A., Knol J. 2016. Early-Life Events, Including

338   Mode of Delivery and Type of Feeding, Siblings and Gender, Shape the Developing Gut

339   Microbiota. *PloS one* 11:e0158498. DOI: 10.1371/journal.pone.0158498.

340 Matsuki T., Watanabe K., Tanaka R., Fukuda M., Oyaizu H. 1999. Distribution of bifidobacterial

341   species in human intestinal microflora examined with 16S rRNA-gene-targeted species-

342   specific primers. *Applied and environmental microbiology* 65:4506–12.

343 Milani C., Mancabelli L., Lugli GA., Duranti S., Turroni F., Ferrario C., Mangifesta M.,

344   Viappiani A., Ferretti P., Gorfer V., Tett A., Segata N., van Sinderen D., Ventura M. 2015.

345   Exploring vertical transmission of bifidobacteria from mother to child. *Applied and*

346   *Environmental Microbiology* 81:7078–7087. DOI: 10.1128/AEM.02037-15.

347 O'Callaghan A., Bottacini F., O'Connell Motherway M., van Sinderen D. 2015. Pangenome

348   analysis of Bifidobacterium longum and site-directed mutagenesis through by-pass of

349   restriction-modification systems. *BMC genomics* 16:832. DOI: 10.1186/s12864-015-1968-4.

350 Roger LC., Costabile A., Holland DT., Hoyles L., McCartney AL. 2010. Examination of faecal

351   Bifidobacterium populations in breast- and formula-fed infants during the first 18 months of

352   life. *Microbiology* 156:3329–3341. DOI: 10.1099/mic.0.043224-0.

353 Ruhaak LR., Stroble C., Underwood MA., Lebrilla CB. 2014. Detection of milk oligosaccharides

354   in plasma of infants. *Analytical and Bioanalytical Chemistry* 406:5775–5784. DOI:

355   10.1007/s00216-014-8025-z.

356 Sela DA., Chapman J., Adeuya A., Kim JH., Chen F., Whitehead TR., Lapidus A., Rokhsar DS.,

357   Lebrilla CB., German JB., Price NP., Richardson PM., Mills DA. 2008. The genome

358   sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization

359   within the infant microbiome. *Proc Natl Acad Sci U S A* 105:18964–18969. DOI:

360   0809584105 [pii]10.1073/pnas.0809584105.

361 Sheu S-J., Hwang W-Z., Chiang Y-C., Lin W-H., Chen H-C., Tsen H-Y. 2010. Use of tuf gene-

362   based primers for the PCR detection of probiotic Bifidobacterium species and enumeration

363   of bifidobacteria in fermented milk by cultural and quantitative real-time PCR methods.

364   *Journal of food science* 75:M521-7. DOI: 10.1111/j.1750-3841.2010.01816.x.

365   Tannock GW., Lawley B., Munro K., Lay C., Taylor C., Daynes C., Baladjay L., Mcleod R.,

366   Thompson-Fawcett M. 2012. Comprehensive analysis of the bacterial content of stool from

367   patients with chronic pouchitis, normal pouches, or familial adenomatous polyposis pouches.

368   *Inflammatory Bowel Diseases* 18:925–934.

369   Turroni F., Peano C., Pass D a., Foroni E., Severgnini M., Claesson MJ., Kerr C., Hourihane J.,

370   Murray D., Fuligni F., Gueimonde M., Margolles A., De Bellis G., O'Toole PW., van

371   Sinderen D., Marchesi JR., Ventura M. 2012. Diversity of bifidobacteria within the infant

372   gut microbiota. *PloS one* 7:e36957. DOI: 10.1371/journal.pone.0036957.

373   Untergasser A., Cutcutache I., Koressaar T., Ye J., Faircloth BC., Remm M., Rozen SG. 2012.

374   Primer3--new capabilities and interfaces. *Nucleic acids research* 40:e115. DOI:

375   10.1093/nar/gks596.

376   Vazquez-Gutierrez P., Lacroix C., Chassard C., Klumpp J., Stevens MJA., Jans C. 2015.

377   Bifidobacterium pseudolongum Strain PV8-2, Isolated from a Stool Sample of an Anemic

378   Kenyan Infant. *Genome announcements* 3:2164. DOI: 10.1128/genomeA.01469-14.

379   Wang RF., Cao WW., Cerniglia CE. 1996. PCR detection and quantitation of predominant

380   anaerobic bacteria in human and animal fecal samples. *Applied and environmental*

381   *microbiology* 62:1242–7.

382   Yatsunenko T., Rey FE., Manary MJ., Trehan I., Dominguez-Bello MG., Contreras M., Magris

383   M., Hidalgo G., Baldassano RN., Anokhin AP., Heath AC., Warner B., Reeder J.,

384   Kuczynski J., Caporaso JG., Lozupone C a., Lauber C., Clemente JC., Knights D., Knight

385   R., Gordon JI. 2012. Human gut microbiome viewed across age and geography. *Nature*

386   486:222–7. DOI: 10.1038/nature11053.

387   Youn SY., Seo JM., Ji GE. 2008. Evaluation of the PCR method for identification of

388   Bifidobacterium species. *Letters in applied microbiology* 46:7–13. DOI: 10.1111/j.1472-

389   765X.2007.02263.x.

390 Young SL., Simon MA., Baird MA., Tannock GW., Bibiloni R., Spencely K., Lane JM.,

391   Fitzharris P., Crane J., Town I., Addo-Yobo E., Murray CS., Woodcock A. 2004.

392   Bifidobacterial Species Differentially Affect Expression of Cell Surface Markers and

393   Cytokines of Dendritic Cells Harvested from Cord Blood. *Clinical and Vaccine*

394   *Immunology* 11:686–690. DOI: 10.1128/CDLI.11.4.686-690.2004.

395 Zhou SJ., Sullivan T., Gibson RA., Lonnerdal B., Prosser CG., Lowry DJ., Makrides M. 2014.

396   Nutritional adequacy of goat milk infant formulas for term infants: a double-blind

397   randomised controlled trial. *The British journal of nutrition* 111:1641–1651. DOI:

398   10.1017/S0007114513004212.

399

400 **Table 1.** PCR primers and probes.

Target	Primer/Probe	Sequence 5' – 3'	Reference
B. longum subsp. *infantis* sialidase gene	inf_2348_F	ATACAGCAGAACCTTGGCCT	This study
	inf_2348_R	GCGATCACATGGACGAGAAC	This study
	inf_2348_P	/FAM/TTTCACGGA/ZEN/TCACCGG ACCATACG/3IABkFQ/	This study
B. longum subsp. *longum* sugar kinase gene	lon_0274_F	GAGGCGATGGTCTGGAAGTT	This study
	lon_0274_R	CCACATCGCCGAGAAGATTC	This study
	lon_0274_P	/56-FAM/AATTCGATG/ZEN/CCCAGCG TGGTCTT/3IABkFQ/	This study
All bacteria *16S rRNA* gene	Uni_F	ACTCCTACGGGAGGCAGCAGT	(Wang, Cao 1996)
	Uni_R	ATTACCGCGGCTGCTGGC	

401
402

403
404 **Table 2.** Effect of DNA extraction method.
405

Sample	Extraction Method	*B. longum* subsp. *longum*[a]	*B. longum* subsp. *infantis*[a]
58550	MoBio	51.75	0.00
58550	Phenol	34.07	0.00
58550	Qiagen	66.85	0.00
AF26B	MoBio	1.66	0.14
AF26B	Phenol	0.18	0.19
AF26B	Qiagen	0.72	0.62
AF12A	MoBio	78.61	0.00
AF12A	Phenol	75.17	0.00
AF12A	Qiagen	75.77	0.00
AF92A	MoBio	0.20	0.00
AF92A	Phenol	0.02	0.00
AF92A	Qiagen	0.06	0.00

406
407 [a]Mean % (duplicates) of total microbiota determined by differential qPCR.

408

409   **Table 3.** Prevalence of *B. longum* subsp. *longum* and subsp. *infantis* in the feces of infants as

410   detected by qPCR.

Geographical location	Nutrition	Delivery	Age at sampling (weeks)	*B. longum* subsp. *longum*[b]	*B. longum* subsp. *infantis*[b]
Chinese[a]	Breast milk	Caesarean	6	86.2	24.1
			8	95.8	12.5
			10	100.0	11.1
			12	82.6	17.4
	Cow's milk formula	Caesarean	6	64.9	27.0
			8	62.2	32.4
			10	81.1	27.0
			12	78.4	32.4
Australian[a]	Breast milk	Vaginal	8	100.0	13.0
		Caesarean	8	85.7	14.3
	Cow's milk formula	Vaginal	8	100.0	4.8
		Caesarean	8	100.0	0.0
	Goat's milk formula	Vaginal	8	100.0	4.8
		Caesarean	8	100.0	0.0
South-East Asian[a]	Breast milk supplemented with cow's milk formula	Vaginal	<1	50.0	8.3
			8	58.3	8.3
			16	58.3	25.0
	Breast milk supplemented with cow's milk formula	Caesarean	<1	21.4	7.1
			8	53.9	15.4
			16	84.6	15.4

411
412   [a]Chinese babies, n = 51 caesarean delivery/breast milk, n = 40 caesarean delivery/formula;

413   Australian babies, n = 30 breast milk (23 vaginal delivery, 7 caesarean), n = 30 cow's milk

414   formula (21 vaginal delivery, 9 caesarean), n = 30 goat's milk formula (21 vaginal, 9 caesarean);

415   South-East Asian babies, n = 12 vaginally delivery/breast milk and/or cow's milk formula, n =

416   12 caesarean delivery/ breast milk and/or cow's milk formula.

417   [b] % of infants harboring the subsp. as determined by differential qPCR.

418

419   **Table 4.** Relative abundances of *B. longum* subsp. *longum* and subsp. *infantis* in the feces of

420   infants as detected by qPCR.

Geographical location	Nutrition	Delivery	Age at sampling (weeks)	*B. longum* subsp. *longum*[b]	*B. longum* subsp. *infantis*[b]
Chinese[a]	Breast milk	Caesarean	6	16.6 (5.4)	2.8 (1.9)
			8	20.4 (6.6)	4.4 (3.6)
			10	26.4 (6.5)	3.6 (3.6)
			12	28.0 (7.3)	1.4 (1.4)
	Cow's milk formula	Caesarean	6	6.6 (3.0)	1.7 (1.3)
			8	11.3 (5.0)	2.3 (1.7)
			10	12.3 (4.6)	4.0 (3.4)
			12	17.7 (6.2)	0.6 (0.3)
Australian[a]	Breast milk	Vaginal	8	33.7 (7.5)	1.1 (1.1)
		Caesarean	8	13.9 (13.4)	2.0 (2.0)
	Cow's milk formula	Vaginal	8	30.1 (7.6)	0.0 (0.0)[c]
		Caesarean	8	13.5 (10.4)	0.0 (0.0)[c]
	Goat's milk formula	Vaginal	8	16.1 (4.5)	0.0 (0.0)[c]
		Caesarean	8	7.2 (4.7)	0.0 (0.0)[c]
South-East Asian[a]	Breast milk supplemented with cow's milk formula	Vaginal	<1	7.3 (3.0)	0.0 (0.0)[c]
			8	4.9 (2.2)	4.6 (4.6)
			16	9.1 (7.5)	1.3 (1.3)
	Breast milk supplemented with cow's milk formula	Caesarean	<1	0.1 (0.1)	0.0 (0.0)[c]
			8	15.1 (4.9)	0.2 (0.2)
			16	18.9 (6.1)	0.0 (0.0)[c]

421   [a]Chinese babies, n = 51 caesarean delivery/breast milk, n = 40 caesarean delivery/formula;

422   Australian babies, n = 30 breast milk (23 vaginal delivery, 7 caesarean), n = 30 cow's milk

423   formula (21 vaginal delivery, 9 caesarean), n = 30 goat's milk formula (21 vaginal, 9 caesarean);

424   South-East Asian babies, n = 12 vaginally delivery/breast milk and/or cow's milk formula, n =

425   12 caesarean delivery/ breast milk and/or cow's milk formula.

426   [b]Mean % (SEM) of total microbiota determined by differential qPCR.

427   [c]Less than 0.01%.

428

429

430

431
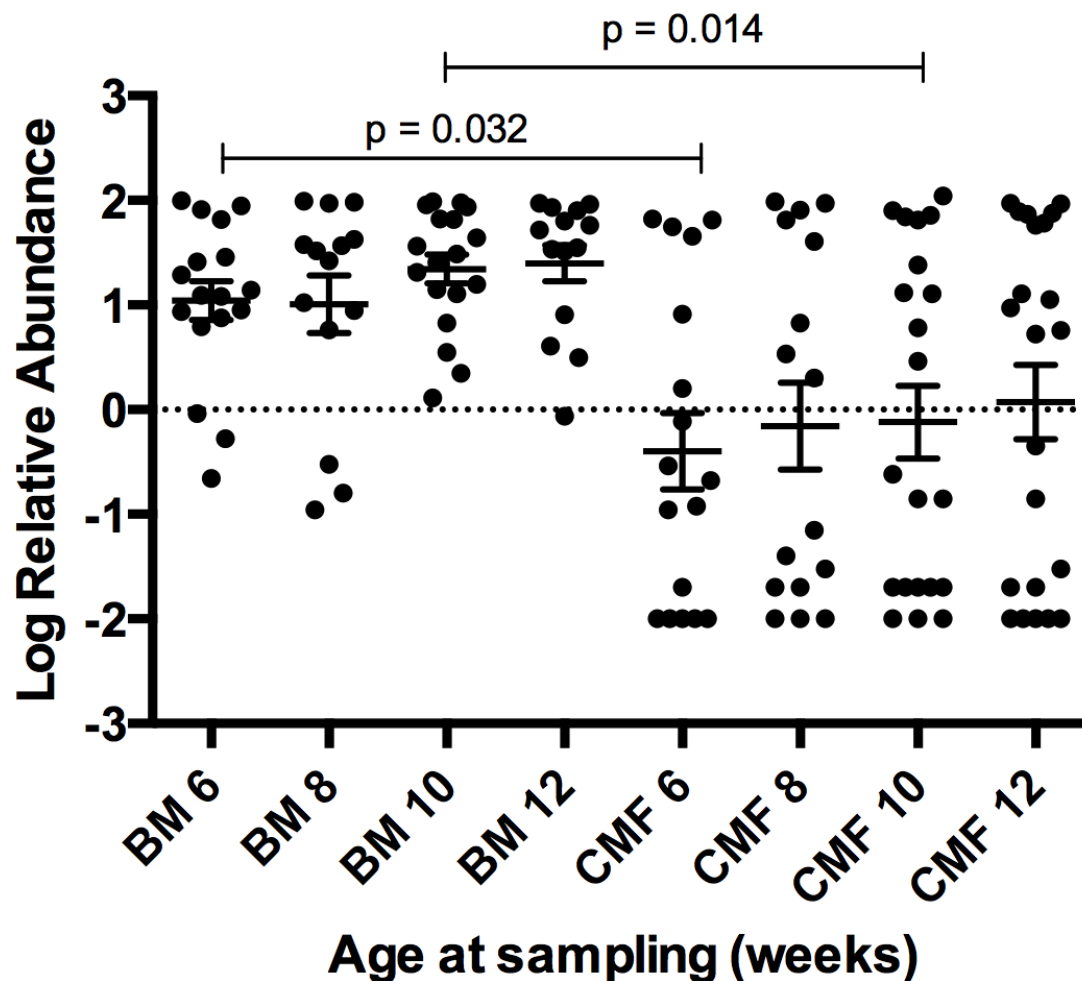
Figure 1. Effect of baby nutrition (breast milk, cow's milk formula) on relative abundances of *B.*

longum subsp. *longum* in the feces of caesarean-delivered Chinese children. The relative

abundance data was normalized by log transformation and evaluated statistically by one-way

ANOVA with multiple comparisons. Scatter plots with means and SEM are shown. Subspecies

longum abundances were more variable in the feces of babies fed cow's milk formula (CMF),

compared to those of breast milk-fed infants (BM) of the same ages. Statistically significant

differences were observed between the 6 and 10 week sampling groups.