 doi:10.20944/preprints202202.0350.v1

Article

Bifidobacterium Species Colonization in Infancy: A Global Cross-Sectional Comparison by Population History of BreastFeeding

Diana H. Taft[1,2], Zachery T. Lewis[1#], Nhu Nguyen[1], Steve Ho[1], Chad Masarweh[1], Vanessa Dunne-Castagna[1], Daniel J. Tancredi[3], M. Nazumul Huda[4,5], Charles B. Stephensen[4,5], Katie Hinde[6], Erika von Mutius[7,8], Pirkka Kirjavainen[9,10], Jean-Charles Dalphin[11ǂ], Roger Lauener[12], Josef Riedler[13], Jennifer T. Smilowitz[1,2], J. Bruce German[1,2], Ardythe L. Morrow[14*] and David A. Mills[1,15*]

[1] Department of Food Science and Technology, University of California Davis, Davis, CA, USA
[2] Foods for Health Institute, University of California Davis, Davis, CA, USA
[3] Department of Pediatrics, University of California Davis School of Medicine, Sacramento, CA, USA
[4] US Department of Agriculture, Western Human Nutrition Research Center, Davis, CA, USA
[5] Department of Nutrition, University of California Davis, Davis, CA, USA
[6] Center for Evolution and Medicine, School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, USA
[7] Dr von Hauner Children's Hospital, Ludwig Maximilian University, Munich, Germany
[8] Institute for Asthma and Allergy Prevention, Helmholtz Centre Munich, Germany
[9] Environment Health Unit, National Institute for Health and Welfare, Kuopio, Finland
[10] Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland
[11] University of Besançon, Department of Respiratory Disease, UMR/CNRS 6249 Chrono-environment, University Hospital, Besançon, France
[12] Christine Kühne-Center for Allergy Research and Education, Davos, Switzerland
[13] Children's Hospital Schwarzach, Austria
[14] Department of Environmental and Public Health Sciences, University of Cincinnati, Cincinnati, OH, USA
[15] Department of Viticulture and Enology, University of California Davis, Davis, CA, USA
* Correspondence: damills@ucdavis.edu(D.A.M.), Tel.: (530)754-7821; morrowa@ucmail.uc.edu(A.L.M.), Tel.: 26 (513)558-0809
#Zachery T. Lewis now employed by Synbiotic Health, Lincoln, NE, USA
ǂDeceased

Abstract: *Bifidobacterium* are a beneficial and dominant member of the breast-fed infant gut microbiome. However, the health benefits of *Bifidobacterium* are partially species dependent. Here we characterize the species and subspecies of *Bifidobacterium* present in breastfed infants around the world. Across populations, three distinct patterns of *Bifidobacterium* colonization emerged: 1) Dominance of *Bifidobacterium longum* subspecies *infantis*, 2) Prevalent *Bifidobacterium* of multiple species, and 3) Frequent absence of *Bifidobacterium.* These patterns appear related to country history of breastfeeding, with infants in countries with historically high rates of long duration breastfeeding more likely to be colonized by *B. longum* subspecies *infantis* compared with infants in countries with histories of shorter duration breastfeeding. These findings highlight the need to consider historical and cultural influences on gut commensal survival influence present day colonization patterns in order to understand epidemiological transmission patterns of *Bifidobacterium* and other major gut commensals.

 doi:10.20944/preprints202202.0350.v1

Keywords: infant; breastfeeding; gut microbiome; Bifidobacterium

1. Introduction

Diet is a major driver of the infant gut microbiome.[1] The earliest studies of the infant gut microbiome were conducted a century ago and highlight the difference in the gut microbiome of breastfed infants and formula fed infants, with breastfed infants exhibiting near monocultures of *Bifidobacterium* and formula fed infants exhibiting a more mixed microbiome without *Bifidobacterium*.[2] The ability of *Bifidobacterium* to dominate the breast-fed infant gut is driven, in part, by human milk oligosaccharides (HMOs),[3-5] the third most abundant solid component of human milk.[6,7] Despite the abundance of HMOs in human milk, these specialized carbohydrates are not digested by the infant.[7] Instead, HMOs act as prebiotics for beneficial bacteria, including *Bifidobacterium*, and as decoy receptors for pathogens.[8] Some *Bifidobacterium* are efficient consumers of HMOs, to the point that HMOs are identified as the "bifidogenic factor" of human milk.[3,9]

Not all *Bifidobacterium* are equally efficient at consuming HMOs,[10] and different species of *Bifidobacterium* can have different implications for infant health.[11] *Bifidobacterium* is a genus composed of 54 different species that had been identified as of January 2017.[12] Some *Bifidobacterium*, such as *B. longum* subsp. *infantis*, are efficient consumers of HMOs, while others, such as *B. longum* subsp. *longum* are better adapted for consumption of plant oligosaccharides.[10] *Bifidobacterium* colonization has a number of potential health benefits for infants, including reduction of allergies,[13,14] improved vaccine response,[15] reduced carriage of antimicrobial resistance genes,[16] reduced carriage of virulence factor genes,[17] reduced enteric inflammation,[18] and immunoregulation via microbial metabolites.[19] However, some of these health benefits are dependent on which species of *Bifidobacterium* are present in the gut microbiome.[15,20] Therefore, it is important to understand variation in patterns of colonization with *Bifidobacterium* in infant populations around the world.

Recent work demonstrates the importance of microbial-accessible carbohydrates (MAC) to maintaining commensal species in host microbiomes.[21] MACs are carbohydrates that are available for fermentation by the microbiome but are not digested by the host. Mouse work has shown that to understand microbes in the present, past dietary practices must also be considered, as populations of mice switched between low and high MAC diets experienced extinctions of portions of the microbes previously present.[21] For breastfed infants, human milk oligosaccharides (HMOs) constitute a major component of milk and function as MACs. Human milk substitutes have typically lacked HMOs, although a few formulas first introduced in 2018 do contain a single HMO. Despite this, HMOs are not typically included in infant formula, and when they are included, the diversity of HMO structures is far below that found in human milk. As a result, infants fed artificial formulas are consuming a diet low in HMOs, and therefore a diet that is low MAC. Because loss of MACs in the diet can drive extinctions in the gut microbiome,[21] populations that experienced periods of low breastfeeding and high formula feeding may have lost species of *Bifidobacterium* most dependent on HMO consumption. In fact, there is already some evidence that this has occurred in

 doi:10.20944/preprints202202.0350.v1

Western populations as *B. longum* subsp. *infantis* has become increasingly difficult to detect in these populations, even in infants never exposed to antibiotics.[22] Therefore, historical breastfeeding patterns must be considered in addition to current breastfeeding in order to understand the relationship between breastfeeding and the infant gut microbiome in the present day. Infants from populations that experienced periods of lower breastfeeding initiation and shorter duration may be living in regions where HMO-consuming *Bifidobacterium* are near extinction due to loss in prior generations.

In particular, we hypothesize that commensals such as *B. longum* subsp. *infantis* dependent on HMO consumption may be lost from populations with a history of lower breastfeeding initiation and duration while *Bifidobacterium* species capable of consuming both HMOs and plant oligosaccharides may have a better chance at surviving in populations that have experienced low breastfeeding rates. Here we work to describe the species of *Bifidobacterium* found in global infant populations, and to consider how population history of breastfeeding practices relates to the currently observed patterns.

2. Methods

Inclusion Criteria for Cohorts

For this analysis, published and unpublished cohorts of term infants where at least some infants were breastfed were selected for inclusion. Cohorts needed to have a minimum of 20 infants and have species-level data on the relative abundance of the *Bifidobacterium* as measured by 16S rRNA gene sequencing and *Bifidobacterium* specific terminal restriction fragment length polymorphism (Bif-TRFLP) and *Bifidobacterium* Longum-Infantis Ratio (BLIR) analysis in infant stool between ages 1 or 2 months, or have 16S rRNA gene sequencing data available with DNA available for species level analysis of *Bifidobacterium*. Publicly available datasets were included from Gambia,[23] Bangladesh,[15] and the PASTURE cohorts from Austria, Finland, Germany, and Switzerland,[24,25] although this is the first publication to include species level *Bifidobacterium* data from the PASTURE cohorts. Additionally, data from a subset of infants enrolled in the University of California Davis Lactation Study (Davis, California)[26,27] and a subset of infants enrolled in the Pediatric Respiratory and Enteric Virus Acquisition and Immunogenesis Longitudinal (PREVAIL) study (Cincinnati, OH)[28] were included. Table 1 summarizes the cohorts, and lists the total number of infants in each cohort, the number of infants who had a sample from ages 1 to 2 months for inclusion in this study, and the number of infants who were at least partially breastfed at time of sample collection. Historically breastfeeding patterns were defined by searching the literature for references on breastfeeding rates published between 1900 and the present day. A high breastfeeding pattern was defined as one where breastfeeding initiation was nearly universal and breastfeeding duration was typically longer than 1 year. The medium breastfeeding pattern was defined as high current and past rates of breastfeeding initiation, but with evidence of historically short (median less than 6 months) duration. The low breastfeeding pattern was defined as occurring in the case of a documented period where at least half of infants were never breastfed (meaning breastfeeding was not initiated), regardless of current population breastfeeding initiation and duration. While breastfeeding rates for each cohort are calculated using all infants with available breastfeeding data in each cohort

 doi:10.20944/preprints202202.0350.v1

(Table 1), only infants who were at least partially breastfed at the time of the month 1 or 2 sample are included in all subsequent analyses in this paper. The UC Davis Lactation Study was approved by the UC Davis Institutional Review Board, and all mothers provided written informed consent for their and their infant's participation in this study. The PREVAIL cohort was approved by the Centers for Disease Control and Prevention (CDC) Institutional Review Board and the Cincinnati Children's Hospital Medical Center Institutional Review Board, and all mothers provided written informed consent for their and their infant's participation in this study. For the PASTURE cohorts, the study was approved by local research ethics committees and written informed consent was obtained from the infant's parents. The Bangladeshi cohort was approved by Research Review Committee (RRC) and Ethical Review Committee (ERC) of International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b). WHO Ethics Review Committee also approved the protocol. Mothers provided written informed consent. For the Gambian cohort, all data were used for this project were publicly available.[23]

For each cohort, data on median duration and initiation rates of breastfeeding are reported if available for the cohort. In addition, historical breastfeeding practices of each country of origin of a cohort are described based on reviews of the published literature. In addition to the month 1-2 samples described above, six cohorts had available longitudinal samples. The Bangladesh cohort had additional samples from a subset of infants at age 2 years with 16S rRNA gene sequencing and Bif-TRFLP/BLIR. The four PASTURE cohorts had additional samples from age 1 year on a subset of infants. The Davis cohort had additional samples from a subset of infants at 3 days of life and age 1 month with 16S rRNA gene sequencing and Bif-TRFLP/BLIR analysis. These cohorts will permit a limited examination of *Bifidobacterium* colonization over time.

16. S rRNA Gene Sequencing and Bif-TRFLP/BLIR

All infant stool samples were extracted and sequenced as described in Davis et al.[23] Analysis of the 16S rRNA gene sequencing results of all raw data files was completed using QIIME2[29] (version qiime2-2017.8) and DADA2.[30] Identification of *Bifidobacterium* species and *B. longum* subsp. was completed using *Bifidobacterium* specific terminal restriction fragment length polymorphism (Bif-TRFLP) and *Bifidobacterium* Longum-Infantis Ratio (BLIR) as described in Davis et al.[23,31] Bif-TRFLP and BLIR are well validated by culture for the identification of *Bifidobacterium* species and subspecies,[23,31] and enable efficient and cost-effective identification of *Bifidobacterium* in a large number of samples.

Relative abundance of total *Bifidobacterium* was compared by cohort and historical breastfeeding pattern using a Kruskal-Wallis test, followed by Dunn's test with Bonferroni correction if results were significant. Prevalence of *Bifidobacterium* species present in at least two cohorts were compared based on breastfeeding history in each cohort using generalized estimating equations (GEE) as implemented in the gee package in R version 3.6.3 with a binomial family and logit linker, cohort of origin used as the clustering variable, and using an exchangeable correlation structure. A *Bifidobacterium* species was considered present in an infant if there was any detectable level of that species in the infant's microbiome. A *Bifidobacterium* species was

 doi:10.20944/preprints202202.0350.v1

considered present in a cohort if that species was present in at least one infant in that cohort. The species included for this analysis were *B. adolescentis, B. animalis, B. bifidum, B. breve, B. longum* subsp. *infantis, B. longum* subsp. *longum, B. longum* of unknown subsp., and *B. pseudocatenulatum.* This meant eight separate models were constructed, leading to a Bonferroni corrected p-value of less than 0.0062 to be considered significant after adjusting for multiple comparisons.

To understand the extent to which individual species of *Bifidobacterium* shape the microbiome, we sought to identify species that, when present in an infant, typically dominated (>50% median relative abundance) or that typically contributed a substantial fraction towards dominance (>20% median relative abundance) in each cohort. To do this, for each species of *Bifidobacterium* in each cohort we selected only infants who were colonized by detectable levels of that species, then calculated the median relative abundance in the colonized infants.

Finally, in the cohorts with samples from infants collected at older ages, Chi-square tests were used to compare the chances of detecting *B. longum* subsp. *infantis* in infants who were still breastfed compared to those who were weaned.

3. Results

Study cohorts and breastfeeding patterns.

This study included eight cohorts with a total of 979 breastfed infants with stool samples collected at age 1 or 2 months. Two cohorts were conducted in two low income countries – Bangladesh and the Gambia - and six cohorts were conducted in five high income countries – Austria, Finland, Germany, Switzerland and the United States (US). The countries included have three historically different breastfeeding patterns, described here as "high", "medium", and "low" breastfeeding.

The first breastfeeding pattern, "high", is characterized by consistent, high rates of breastfeeding initiation and long duration of breastfeeding without evidence of historical interruptions in either breastfeeding initiation or duration. This is the pattern observed in The Gambia and Bangladesh. In the 1980s, when breastfeeding was low in much of the world, 97.5% of Bangladeshi children were fed at least some breastmilk and the mean duration of breastfeeding was 26 months [32]. Furthermore, this pattern of high rates of breastfeeding and long duration remained consistent in studies of Bangladesh from the 1970s through the 1990s.[33] More recently, the initiation of breastfeeding in Bangladesh remains high (98.3%) and duration of breastfeeding has increased to a mean of 31.9 months.[34] The Gambia has a similar historical and modern pattern, where a report from 1979 points out that "all Gambian women breast feed their babies for up to two years."[35] Despite a shorter than optimal duration of exclusive breastfeeding, breastfeeding rates remain high in the Gambia today, with a median duration of 20 months with over 95% of infants ever breastfed.[36] Therefore, in this study, infants in the Bangladesh and Gambian cohorts were therefore considered to have a "high" historical breastfeeding pattern. Median duration of breastfeeding is unknown for the Gambian cohort, but all infants in this group were breastfed for at least 5 months as inclusion criteria included the availability of week 20 milk samples.[23] Median duration of breastfeeding in the Bangladeshi cohort

 doi:10.20944/preprints202202.0350.v1

is also unknown as infants in this cohort were only followed until age 2 years, and more than half of the infants enrolled in the study were still breastfed at this time point (Table 2). This means that the high breastfeeding pattern is still observed in the present day in these countries.

The second breastfeeding pattern, "medium", is characterized by high, consistent breastfeeding initiation but historical evidence of short duration of breastfeeding (less than 6 months of any breastfeeding in the majority of infants). This pattern is observed in Austria, Finland, Germany, and Switzerland, with historical documentation of a median duration of breastfeeding of less than 6 months during the 1970s and 1980s. In European countries, there was a general decline in breastfeeding following World War II followed by increased rates of breastfeeding starting in the 1970s.[37] In Austria, Germany, and Finland, in the 1970s and 1980s most women initiated breastfeeding, but duration was short.[38] In Austria, only 5% of infants were breastfed at 3 months post-partum in 1980, which increased to 41% by 1984,[39] indicating that the majority of infants did not receive a full six months of breastfeeding during the early 1980s. In Finland in the early 1980s, most women initiated breastfeeding but only a third of infants were breastfed until 3 months,[40] again indicating a historical period where the majority of infants were not at least partially breastfed for six months. In Germany, only 2% of infants were still breastfed at age 6 months in 1980.[41] In Switzerland, in 1978, 92% of mothers initiated breastfeeding, however only 30% were still breastfeeding by 4 months.[42] These past patterns are no longer present in these countries. In Austria today, breastfeeding initiation rates are currently about 98% with a median duration of 27 weeks.[43] In Finland today, almost all mothers initiate breastfeeding, 60% are still breastfeeding at 6 months and a third are still breastfed at 11 months of age.[44] In Germany, 90% of women initiate breastfeeding and more than half of all infants are still breastfed at 6 months.[45] In Switzerland, breastfeeding initiation is 95% and the median duration of any breastfeeding is 31 weeks.[46] These numbers are consistent with those observed for the study cohorts (Table 2). Therefore, even though present day practices in these countries would qualify for the "high" category, historical evidence supports a period with low HMO consumption by infants creating the opportunity for shifts in infant *Bifidobacterium* populations.

The third pattern of breastfeeding, "low", is characterized by low historic rates of breastfeeding initiation. The two cohorts from the United States meet this pattern, as the United States experienced a low point in breastfeeding in 1971 when less than 1 in 4 women breastfed their infants even once.[47] Following sustained public health efforts, rates partially recovered to approximately 50%-60% initiating breastfeeding in the 1980s.[48] Breastfeeding initiation remained at only 60% in 1995, with only 22% of mothers still breastfeeding at age 6 months in 1995.[49] As recently as 2004, the United States had lower breastfeeding initiation and duration rates than many continental European countries.[50] Since that time, breastfeeding in the US has increased so that in 2017, 84% of infants were ever breastfed and 58% were still breastfed at 6 months.[51] In the state of Ohio (where Cincinnati is located) in 2017 (the year that the Cincinnati PREVAIL cohort began enrollment), 80% of infants were ever breastfed and 51% were still breastfed at age 6 months.[51] For the state of California (where Davis is located) in 2009 (the year the UC Davis Lactation Cohort began enrollment), 85% of infants were ever breastfed and 53% were

 doi:10.20944/preprints202202.0350.v1

still breastfed at age 6 months.[52] Because the Davis Lactation Cohort enrolled based on the intention to breastfed, this cohort has a longer median duration of breastfeeding than was observed in the general population (Table 2). The Cincinnati cohort had a lower duration of breastfeeding (Table 2) than that reported for Ohio, in part because this cohort is an urban, diverse cohort which tends to mean a shorter duration of breastfeeding.

Bifidobacterium species across Cohorts and Breastfeeding patterns.

Average *Bifidobacterium* levels differed between cohorts (Figure 1). The prevalence of detection of any *Bifidobacterium* was 100% of infants in all cohorts except those in the United States. The total relative abundance of *Bifidobacterium* in individual infants also differed significantly by cohort (Kruskal-Wallis test, $p<0.0001$) and by breastfeeding pattern (Kruskal-Wallis test, $p<0.0001$). By cohort, Davis had significantly lower total relative abundance of *Bifidobacterium* than all other cohorts (Dunn's test, $p<0.0001$ compared to all cohorts except Cincinnati where $p=0.03$). Switzerland had significantly higher relative abundance of *Bifidobacterium* than all other cohorts except for Gambia (Dunn's test, $p<0.0001$ for all cohorts except Gambia, no significant difference in total *Bifidobacterium* between Switzerland and Gambia, $p=1.0$). There were no other significant differences in total *Bifidobacterium* relative abundance by cohort. By historical breastfeeding status, the low breastfeeding pattern infants had a lower relative abundance of *Bifidobacterium* than the high breastfeeding pattern infants (Dunn's test, $p<0.0001$) and the medium breastfeeding pattern infants (Dunn's test, $p<0.0001$). The medium breastfeeding pattern infants had higher total *Bifidobacterium* than the high breastfeeding infants (Dunn's test, $p=0.02$).

Importantly, the cohorts differed in the species of *Bifidobacterium* that colonized infant guts. The average infant gut microbiome in Bangladesh and Gambia is dominated (meaning greater than 50% relative abundance *Bifidobacterium* of a single subspecies) by *B. longum* subsp. *infantis,* with an average relative abundance of 54% *B. longum* subsp. *infantis* in Bangladesh and 53% in the Gambia. In the European and USA cohorts, the average infant gut microbiome was not dominated by any single species of *Bifidobacterium* (Figure 1).

The presence or absence of particular *Bifidobacterium* species in infants reflected historical breastfeeding practices among the cohorts (see Table 3 for the summary of findings from GEE models). Notably, *B. longum* subsp. *infantis* and *B. bifidum* were significantly less likely to be present in infants from medium and low breastfeeding pattern cohorts compared to infants in high breastfeeding pattern cohorts. *B. pseudocatenulatum* was more likely to be present in infants from medium or low breastfeeding pattern cohorts than in infants from high breastfeeding pattern cohorts. *B. adolescentis* and *B. breve* were detected more often in medium breastfeeding pattern cohorts than in high breastfeeding pattern cohorts, but there was no significant difference in the presence of these species between low breastfeeding pattern cohorts and high breastfeeding pattern cohorts. The prevalence of these species in each cohort is shown in Table 2. Because *B. breve*, *B. pseudocatenulatum,* and *B. bifidum* may be difficult to distinguish when using Bif-TRFLP, the values in Table 2 may underestimate the true prevalence of these taxa in all cohorts. All figures include the relative abundance of these mixed peaks as "Unknown

 doi:10.20944/preprints202202.0350.v1

Bifidobacterium", but indistinguishable peaks do not count towards the estimates of relative abundance for these species or towards the presence or absence of these species in any infant.

Detailed visualizations of *Bifidobacterium* colonization of individual infants in each cohort demonstrate the substantial concordance between historical breastfeeding patterns and infant gut colonization (Figure 2). The similarities of infant *Bifidobacterium* colonization patterns by historical breastfeeding are striking. A majority of infants in the high historical breastfeeding cohorts having gut microbiomes dominated by *B. longum* subsp. *infantis* with lower relative abundance of other *Bifidobacterium* species (Figure 2A and 2B). In the medium historical breastfeeding cohorts, infants still have very high levels of total *Bifidobacterium*, but no single species dominates (Figure 2C-2F). In the low historical breastfeeding cohorts, some breastfed infants completely lack *Bifidobacterium,* but those who are colonized by *Bifidobacterium* generally appear similar to the pattern observed in the medium historical breastfeeding cohorts. As medium breastfeeding pattern infants had higher total *Bifidobacterium* levels than high breastfeeding pattern infants and because there were significant differences in the presence and absence of specific *Bifidobacterium* species, we examined which *Bifidobacterium* could dominate the infant gut microbiome. In the high historical breastfeeding pattern cohorts, approximately two-thirds of infants had a gut microbiome dominated (meaning greater than 50% relative abundance from a single source) by *B. longum* subsp. *infantis*. In addition, *B. breve, B. longum* subsp. *longum,* and *B. pseudocatenulatum* or a mixture of multiple *Bifidobacterium* species would occasionally result in an infant gut microbiome dominated by *Bifidobacterium.* In the medium historical breastfeeding pattern cohorts, a microbiome dominated by *B. longum* subsp. *infantis* was rare, but roughly two-thirds of infants had a gut microbiome dominated by *B. breve, B. longum* subsp. *longum,* and *B. pseudocatenulatum* or a mixture of multiple *Bifidobacterium* species. In the low historical breastfeeding pattern cohorts, some breastfed infants completely lacked any detectable level of *Bifidobacterium.* When infants did have gut microbiomes dominated by *Bifidobacterium*, it was typically the same mix of species as observed in the medium historical breastfeeding pattern cohorts. On occasion, *B. bifidum* could also dominate an infant gut microbiome, but this was a rare occurrence with only 1 infant in the German cohort, 1 infant in the Finnish cohort, and 1 infant in the Davis, USA cohort.

Timing and Bifidobacterium Species Colonization

Six cohorts, the Bangladesh cohort, the four PASTURE cohorts (from Austria, Finland, Germany, and Switzerland), and the UC Davis lactation cohort, had additional longitudinal data on infant *Bifidobacterium* colonization patterns, although the timing of the available data differed. We analyzed these additional time points to address questions of timing regarding the acquisition of *Bifidobacterium* species (in the UC Davis cohort) and the loss or consistency of colonization with *Bifidobacterium* species comparing early infancy to 1 or 2 years of age (Bangladesh cohort and PASTURE cohorts).

There were differences over time in the colonization of Davis infants with *B. longum* subsp. *infantis* (Figure 3). Figure 3A again shows the 2 month *Bifidobacterium* colonization of the Davis infants, the stars indicate the infants colonized by *B. infantis* who also had samples at an earlier time point.

 doi:10.20944/preprints202202.0350.v1

Regarding timing of acquisition of *Bifidobacterium* species, the UC Davis lactation cohort had additional samples with known *Bifidobacterium* colonization patterns from day 3 of life (29 infants; Figure 3B) and month 1 of life (30 infants; Figure 3C) in addition to the month 2 samples described above. The median relative abundance of total *Bifidobacterium* was only 1% at day 3 (range 0% - 91%), with 38% of infants having no detectable level of *Bifidobacterium* present. The species of *Bifidobacterium* detected at day of life 3 were *B. longum* subsp. *longum* (detected in 28% of infants), *B. bifidum* (in 14% of infants), *B. breve* (in 14% of infants), *B. pseudocatenulatum* (in 14% of infants), and *B. adolescentis* (in 3% of infants) (Figure 4A). At age one month, 30 of the infants (including all 29 with day of life 3 samples) had additional Bif-TRFLP/BLIR results (Figure 4B). 43% of Davis infants had no detectable *Bifidobacterium* at month 1 of life, and a median relative abundance of total *Bifidobacterium* of 0.2% (range 0% - 95%). The species that were present were *B. longum* subsp. *longum* (detected in 23% of infants), *B. bifidum* (in 23% of infants), *B. breve* (in 23% of infants), *B. adolescentis* (in 10% of infants), and *B. pseudocatenulatum* (in 7% of infants). Notably, *B. longum* subsp. *infantis* was not found in any Davis infants prior to age 2 months.

One hundred and nine of the 274 Bangladeshi infants had data on *Bifidobacterium* colonization at age 2 years. At age 2, all of these infants were still colonized with at least some *Bifidobacterium*, but the median total relative abundance of *Bifidobacterium* was lower than in early life at 21% relative abundance *Bifidobacterium* (range 3% - 82%; Figure 4). Furthermore, the prevalence of *B. longum* subsp. *infantis* had dropped considerably and was found in just 17% of infants overall, including 20% of still breastfed infants and 11% of infants who had not received breastmilk in the past week (Table 4). In contrast to the dominance of *B. longum* subsp. *infantis* in early infancy, the species of *Bifidobacterium* found in the Bangladeshi infant at age 2 years, in order of decreasing prevalence, were *B. pseudocatenulatum* (detected in 95% of infants), *B. longum* subsp. *longum* (detected in 95% of infants), *B. bifidum* (detected in 78% of infants), *B. breve* (detected in 65% of infants), *B. adolescentis* (detected in 32% of infants), *B. longum* subsp. *infantis* (detected in 17% of infants), *B. animalis* (detected in 9% of infants), *B. magnum* (detected in 2% of infants), and an indistinguishable mix of *B. choerinum* and/or *B. pseudolongum* (1% of infants, this indistinguishable mix was also plotted as part "Unknown *Bifidobacterium*").

In the Austrian cohort, 91 of the infants breastfed at 2 months also had a year 1 sample. In the Finnish cohort, 129 of the infants breastfed at 2 months also had a year 1 sample. In the German cohort, 110 of the infants breastfed at 2 months also had a 1 year sample. In the Swiss cohort, 175 of the infants breastfed at 2 months also had 1 year sample. All infants were still colonized with at least some *Bifidobacterium* (Figure 5), but usually at a lower relative abundance of total *Bifidobacterium* than was present at 2 months of age. Comparing colonization with *B. longum* subsp. *infantis* between infants who were or were not still breastfed at age 1 year, infants who were breastfed always had a higher prevalence of *B. longum* subsp. *infantis* than those who were not, but this difference only reached significance in the Austrian (Chi-square test, $p=0.003$) and Swiss (Chi-square-test, $p<0.001$) cohorts (Table 4). The greater prevalence of *B. longum* subsp. *infantis* in breastfed at 1 year infants compared to not breastfed at 1 year infants in the Austrian and Swiss cohorts is

 doi:10.20944/preprints202202.0350.v1

consistent with horizontal transmission of *B. longum* subsp. *infantis*, as this species does not remain at high rates in infants who are no longer breastfed. Furthermore, across all the European cohorts, there were 28 infants with detectable levels of *B. longum* subsp. *infantis* at month 2 but not at year 1, 18 infants with detectable levels of *B. longum* subsp. *infantis* at year 1 but not at month 2, and 9 infants with detectable levels at both time points. Of the 18 infants who acquired *B. longum* subsp. *infantis* between 2 months and 1 year, 13 (72%) were breastfed at 1 year of age while only 5 (28%) were not breastfed at 1 year of age. In contrast, of the 28 infants who lost *B. longum* subsp. *infantis* between month 2 and year 1, 26 (93%) were not breastfed at age 1 year and only 2 (7%) were breastfed at age 1 year. This is consistent with the idea that infants who are breastfed are more likely to acquire *B. longum* subsp. *infantis* and that this subspecies is likely to be lost after the end of breastfeeding. The mix of *Bifidobacterium* species found in these infants at 1 year of age was generally similar to those observed at 2 months (Table 5).

4. Discussion

In our study of cohorts from across the world, we found that the infants from high historical breastfeeding pattern countries are more likely to be colonized with *B. longum* subsp. *infantis* and *B. bifidum* than infants from other cohorts. The pattern of *Bifidobacterium* colonization by cohort is consistent with the apparent loss of *B. longum* subsp. *infantis* in locations of the world with historically lower breastfeeding rates and shorter breastfeeding durations. In individual infants, several *Bifidobacterium* can at least occasionally dominate the gut microbiome. However, in this study only *B. longum* subspecies *infantis* dominated the gut microbiome in more than half of the infants where it is highly prevalent. In the relative absence of *B. longum* subsp. *infantis,* the infants from medium and low breastfeeding pattern countries, three species are likely to contribute substantially to the dominance of total *Bifidobacterium* including *B. breve, B. longum* subsp. *longum*, and *B. pseudocatenulatum.* Two of these species, *B. longum* subsp. *longum* and *B. pseudocatenulatum,* are also more likely to colonize infants from medium and low breastfeeding countries. This suggests that these species may contribute more to the infant gut microbiome when species tied tightly to HMO consumption are absent. Infants from medium breastfeeding pattern countries also had higher rates of colonization with *B. adolescentis* and *B. breve* than high breastfeeding pattern countries. This loss of *B. longum* subsp. *infantis* in medium and low breastfeeding countries opens an ecological niche for other *Bifidobacterium,* creating a situation where infants from medium breastfeeding countries frequently have the same relative abundance of *Bifidobacterium* as infants from high breastfeeding pattern countries, but the *Bifidobacterium* are from multiple species rather than dominated by just the single subspecies as occurs with *B. longum* subsp. *infantis.* The infants from low breastfeeding pattern cohorts sometimes lack detectable *Bifidobacterium*, but otherwise appear similar to the infants of medium breastfeeding cohorts. Despite its rarity, when *B. longum* subsp. *infantis* is detected in infants from the medium and low breastfeeding pattern cohorts, it typically dominates the infant gut microbiome of those infants similar to the pattern observed in the high breastfeeding country infants.

 doi:10.20944/preprints202202.0350.v1

The timing of colonization with *B. longum* subsp. *infantis* is distinct from the timing of colonization that occurs with other *Bifidobacterium.* Most species of *Bifidobacterium* appeared early in the Davis cohort, suggesting a pattern of vertical transmission consistent with prior work on the transmission of *Bifidobacterium*.[53] *B. longum* subsp. *Infantis*, however, did not appear as a part of any infant's gut microbiome until age 2 months, suggesting that this particular subspecies may not transmit vertically from the mother. The longitudinal data from the Bangladeshi cohort further supports our hypothesis, as *B. longum* subsp. *infantis* is found in 84% of infants at age 6 weeks, but only 17% of infants at age 2 years. This suggests that even in populations where *B. longum* subsp. *infantis* is commonly found in infants, this subspecies is unlikely to persist in adults as it is replaced by a mix of species similar to those seen in infants from other populations as breastmilk becomes a smaller portion of the diet and the gut microbiome matures. As only 17% of 2 year olds are colonized by *B. longum* subsp. *infantis* it is unlikely that 84% of women are colonized with *B. longum* subsp. *infantis* to transmit it vertically. This is also consistent with a horizontal transmission of *B. longum* subsp. *infantis* between infants. The results from the European cohort also support this trend, as the infants who are breastfed are more likely to be colonized by *B. longum* subsp. *infantis* than those who are not, and a number of infants who did not have detectable levels of *B. longum* subsp. *infantis* at 2 months of age had detectable levels at age 1 year if they remained breastfed.

The pattern of increased prevalence of *B. longum* subsp. *infantis* in cohorts with an uninterrupted history of long duration breastfeeding is consistent with other studies. Consider a population in the US that never widely adopted formula: Old Order Mennonites. Old Order Mennonites are unusual in the United States because they have never widely adopted formula use, and generally avoid antibiotic use.[54] Consistent with what is discussed in the cohorts included in this study, colonization by *B. longum* subsp. *infantis* or with any *Bifidobacterium* is more prevalent in the Old Order Mennonites, with *B. longum* subsp. *infantis* detected by qPCR in 70% of the Old Order Mennonite infants.[13] As *B. longum* subsp. *infantis* is genetically equipped to consume HMOs[10] and formula lacks HMOs, this supports the concept of microbial accessible carbohydrates (MACs) or their lack driving extinctions of microbial commensals within populations. The low MAC diet associated with historical formula use may explain the disappearance of *B. longum* subsp. *infantis* from European and United States cohorts. The higher breastfeeding initiation rates in the European countries may have permitted some species of *Bifidobacterium* to continue to circulate at higher levels in infants than what was possible in the US, particularly if the species in question are also found in the guts of adults and older children. Many adult strains of *Bifidobacterium* have at least some of the genetic machinery needed to consume HMOs[10] and may therefore move in to take advantage of a niche left open by the missing *B. longum* subsp. *infantis*.

This study does have limitations, including the small number of cohorts included. Importantly, the limited number of cohorts is due, in part, to selecting only cohorts of infants with species level identification of *Bifidobacterium* completed using the same technique during a narrow infant age window. This careful limitation to a single method permits direct comparisons across cohorts that would not be possible if a broader range of time points or

 doi:10.20944/preprints202202.0350.v1

methods were included. The small number of included cohorts is balanced by the more than 900 breastfed infants included in this study. An additional strength of this study is the inclusion of cohorts from four different continents and a clear gradient in breast-feeding history across cohorts. Using Bif-TRFLP/BLIR also presents a challenge, where in some cases species identification may be ambiguous. This imprecision in species identification is particularly common with *B. bifidum, B. breve,* and *B. pseudocatenulatum* as these three species produce peaks of very similar size. Most of the unknown *Bifidobacterium* reported in the plots is related to ambiguous peaks generated by some strains of *B. bifidum*, *B. breve,* and *B. pseudocatenulatum*. This suggests that abundance and prevalence estimates for these three species may be lower than what is reported here. *B. bifidum* occurred more frequently in Bangladesh than in any other cohort, but was completely absent in the Gambia. As the number of infants in the Gambian cohort was small, this may represent a lower prevalence of *B. bifidum*, perhaps similar to what is seen in the European countries rather than true absence. Another possibility is that *B. bifidum* is present in this cohort, but undetected because of the similarity of *B. bifidum* to *B. pseudocatenulatum* and *B. breve* when using Bif-TRFLP. Because of these limits in the data, it is difficult to draw conclusions about *B. bifidum* global colonization patterns from this work. The limitation in separating out *B. bifidum, B. breve,* and *B. pseudocatenulatum* is counter-balanced by the improved ability to distinguish *B. longum* subsp. *infantis* from *B. longum* subsp. *longum,* even in the event that both subspecies are found in the same infant with one at much lower levels than the other. As *B. longum* subsp. *infantis* has been proposed to be endangered in western countries,[55] a method that could reliably and efficiently distinguish these two subspecies was critical to this work on how breastfeeding fits into the concept of microbial accessible carbohydrates. An additional limitation of this work is the narrow focus on breastfeeding; there are several additional factors not considered in this work that may be important for *Bifidobacterium* colonization patterns. These factors include prevalence of antibiotic use in the different cohorts, infant care patterns (e.g. when infants are sent to out of home childcare[56]), and hygiene practices. These metadata were not consistently collected for the included cohorts so they could not be considered in the present study. Future work is needed to understand the complex set of factors that support *Bifidobacterium* transmission both historically and in the present day. For the present study, we were less concerned about antibiotic use than diet, because antibiotic use is prevalent in Bangladesh, where the prevalence of *B. longum* subsp. *infantis* remains high. For instance, approximately 40% of children under the age of 5 with an acute respiratory infection in Bangladesh receive antibiotics.[57] This suggests that prevalent antibiotic use is less likely to be a driving factor in the missing *B. longum* subsp. *infantis* than diet.

Bifidobacterium are important infant commensals, but the full health implications of colonization by different species of *Bifidobacterium* remains unclear. This is despite evidence that colonization by *Bifidobacterium* is important to health and evidence that some *Bifidobacterium* species may be more effective at supporting infant health than others. Some benefits of high levels of *Bifidobacterium* colonization do not appear to be species dependent, such as reduced carriage of antimicrobial resistance genes.[16] Other benefits of high levels of *Bifidobacterium* are species dependent. For example, higher levels of

B. longum subsp. *infantis* are associated with improved vaccine response, but the same association is not seen with higher levels of *B. longum* subsp. *longum* or *B. breve*.[15] *B. longum* subsp. *longum* is found more often in healthy infants than in those with allergic symptoms, but the same trend was not seen for other *Bifidobacterium* species.[58] *B. longum* subsp. *infantis* was found more frequently in Old Order Mennonite infants with lower risk of atopic disease than in Rochester infants.[13] As such, understanding the differences in *Bifidobacterium* species colonization in infant populations is important when studying the health implications of the early life microbiome. The fact that the health implications of *Bifidobacterium* vary by species and that *Bifidobacterium* species colonization patterns vary by country and population history of breastfeeding mean care needs to be taken in interpreting health findings associated with of high levels of *Bifidobacterium* without specifying species if trying to apply the findings to other populations. This means it is critical to understand the epidemiological factors that support transmission of different species of *Bifidobacterium*, including considering how historic formula-driven extinctions may have changed the colonization patterns observed in the present day.

Author Contributions: DHT, ZTL, NN, SH, CM, VDC, MNH, CBS, EvM, PK, JCD, RL, JR, JTS, JBG, ALM, and DAM all contributed to the acquisition of data. DHT, ZTL, DJT, ALM, and DAM contributed to data analysis. DHT, ZTL, KH, ALM, and DAM contributed to the conception and design of this study. DHT, ZTL, ALM, and DAM drafted the initial manuscript. All authors critically reviewed and approved the final manuscript, except JCD who passed away before the manuscript draft was complete.

Funding: Funding for this project includes US Department of Agriculture – Agricultural Research Service project 5306-51530-018-00, World Health Organization project 2010168947, Thrasher Research fund grant 11488. Funding for this project also includes National Institutes of Health award F32HD093185 (D.H.T.) and the Peter J. Shields Endowed Chair in Dairy Food Science (D.A.M.).

Informed Consent Statement: The Bangladesh, Gambian, and PASTURE cohorts include already published data. The UC Davis Lactation Study was approved by the UC Davis Institutional Review Board, and all mothers provided written informed consent for their and their infant's participation in this study. The PREVAIL cohort was approved by the Centers for Disease Control and Prevention (CDC) Institutional Review Board and the Cincinnati Children's Hospital Medical Center Institutional Review Board, and all mothers provided written informed consent for their and their infant's participation in this study.

Impact Statement: Population history of breastfeeding influences the present-day infant microbiome. *Bifidobacterium longum* subspecies *infantis*, an efficient consumer of HMOs, is found most commonly in populations with uninterrupted high rates of breastfeeding initiation and uninterrupted long duration of breastfeeding. Other species of *Bifidobacterium* are more commonly found in populations with more disrupted breastfeeding histories.

Acknowledgments: Thank you to all the mother-infant dyads who participated in this research. Funding for this project includes US Department of Agriculture – Agricultural Research Service project 5306-51530-018-00, World Health Organization project 2010168947, Thrasher Research fund grant 11488. Funding for this project also includes National Institutes of Health award F32HD093185 (D.H.T.) and the Peter J. Shields Endowed Chair in Dairy Food Science (D.A.M.).

Conflicts of Interest: David A. Mills and J. Bruce German are co-founders of Evolve Biosystems and BCD Biosciences. Jennifer T. Smilowitz consults with Evolve Biosystems. Evolve Biosystems and BCD Biosciences had no role in the funding, design, data acquisition and analysis, or decision to publish this work. Ardythe L. Morrow is a co-founder of Glycosyn, LLC, which had no role in the funding, design, data acquisition and analysis, or decision to publish this work.

References

1 Davis, E. C., Wang, M. & Donovan, S. M. The Role of Early Life Nutrition in the Establishment of Gastrointestinal Microbial Composition and Function. *Gut microbes* **8**, 143-171 (2017).

2 Logan, W. R. The Intestinal Flora of Infants and Young Children. *The Journal of Pathology and Bacteriology* **18**, 527-551 (1913).

3 Ward, R. E., Ninonuevo, M., Mills, D. A., Lebrilla, C. B. & German, J. B. In Vitro Fermentability of Human Milk Oligosaccharides by Several Strains of Bifidobacteria. *Molecular nutrition & food research* **51**, 1398-1405 (2007).

4 LoCascio, R. G. et al. Glycoprofiling of Bifidobacterial Consumption of Human Milk Oligosaccharides Demonstrates Strain Specific, Preferential Consumption of Small Chain Glycans Secreted in Early Human Lactation. *Journal of Agricultural and Food Chemistry* **55**, 8914-8919 (2007).

5 Sela, D. A. et al. The Genome Sequence of ≪Em≫Bifidobacterium Longum≪/Em≫ Subsp. ≪Em≫Infantis≪/Em≫ Reveals Adaptations for Milk Utilization within the Infant Microbiome. *Proceedings of the National Academy of Sciences* **105**, 18964 (2008).

6 Newburg, D., Neubauer, S. & Jensen, R. Handbook of Milk Composition. *ed Jensen RG, Academic Press, San Diego*, 273-349 (1995).

7 Petherick, A. Development: Mother'S Milk: A Rich Opportunity. *Nature* **468**, S5 (2010).

8 Ballard, O. & Morrow, A. L. Human Milk Composition: Nutrients and Bioactive Factors. *Pediatric Clinics* **60**, 49-74 (2013).

9 Bode, L. Human Milk Oligosaccharides: Every Baby Needs a Sugar Mama. *Glycobiology* **22**, 1147-1162 (2012).

10 Sela, D. A. & Mills, D. A. Nursing Our Microbiota: Molecular Linkages between Bifidobacteria and Milk Oligosaccharides. *Trends in microbiology* **18**, 298-307 (2010).

11 Tannock, G. W. Commentary: Remembrance of Microbes Past. *International Journal of Epidemiology* **34**, 13-15 (2005).

12 Mattarelli, P. & Biavati, B. in *The Bifidobacteria and Related Organisms: Biology, Taxonomy, and Applications* (Mattarelli, P., Biavati, B., Holzapfel, W. H. & wood, B. J. B. eds.) Ch. 2, (Andre G. Wolff, 2018).

13 Seppo, A. E. et al. Infant Gut Microbiome Is Enriched with Bifidobacterium Longum Ssp. Infantis in Old Order Mennonites with Traditional Farming Lifestyle. *Allergy* **n/a** (2021).

14 Cukrowska, B., Bierła, J. B., Zakrzewska, M., Klukowski, M. & Maciorkowska, E. The Relationship between the Infant Gut Microbiota and Allergy. The Role of Bifidobacterium Breve and Prebiotic Oligosaccharides in the Activation of Anti-Allergic Mechanisms in Early Life. *Nutrients* **12**, 946 (2020).

15 Huda, M. N. et al. (Am Assoc Immnol, 2017).

16 Taft, D. H. et al. Bifidobacterial Dominance of the Gut in Early Life and Acquisition of Antimicrobial Resistance. *mSphere* **3** (2018).

17 Casaburi, G. & Frese, S. A. Colonization of Breastfed Infants by Bifidobacterium Longum Subsp. Infantis Evc001 Reduces Virulence Gene Abundance. *Human Microbiome Journal* **9**, 7-10 (2018).

18 Henrick, B. M. et al. Colonization by B. Infantis Evc001 Modulates Enteric Inflammation in Exclusively Breastfed Infants. *Pediatric Research* **86**, 749-757 (2019).

19 Henrick, B. M. et al. Bifidobacteria-Mediated Immune System Imprinting Early in Life. *Cell* **184**, 3884-3898.e3811 (2021).

20 He, F. et al. Comparison of Mucosal Adhesion and Species Identification of Bifidobacteria Isolated from Healthy and Allergic Infants. *FEMS immunology and medical microbiology* **30**, 43-47 (2001).

21 Sonnenburg, E. D. et al. Diet-Induced Extinctions in the Gut Microbiota Compound over Generations. *Nature* **529**, 212-215 (2016).

22 Tannock, G. W., Lee, P. S., Wong, K. H. & Lawley, B. Why Don't All Infants Have Bifidobacteria in Their Stool? *Frontiers in microbiology* **7**, 834 (2016).

23 Davis, J. C. C. et al. Growth and Morbidity of Gambian Infants Are Influenced by Maternal Milk Oligosaccharides and Infant Gut Microbiota. *Scientific Reports* **7**, 40466 (2017).

24 von Mutius, E. & Schmid, S. The Pasture Project: Eu Support for the Improvement of Knowledge About Risk Factors and Preventive Factors for Atopy in Europe. *Allergy* **61**, 407-413 (2006).

25 Depner, M. et al. Maturation of the Gut Microbiome During the First Year of Life Contributes to the Protective Farm Effect on Childhood Asthma. *Nature Medicine* **26**, 1766-1775 (2020).
26 Smilowitz, J. T. et al. The Human Milk Metabolome Reveals Diverse Oligosaccharide Profiles. *The Journal of Nutrition* **143**, 1709-1718 (2013).
27 Lewis, Z. T. et al. Maternal Fucosyltransferase 2 Status Affects the Gut Bifidobacterial Communities of Breastfed Infants. *Microbiome* **3**, 13 (2015).
28 Morrow, A. L. et al. Pediatric Respiratory and Enteric Virus Acquisition and Immunogenesis in Us Mothers and Children Aged 0-2: Prevail Cohort Study. *JMIR research protocols* **10**, e22222 (2021).
29 Caporaso, J. G. et al. Qiime Allows Analysis of High-Throughput Community Sequencing Data. *Nature Methods* **7** (2010).
30 Callahan, B. J. et al. Dada2: High-Resolution Sample Inference from Illumina Amplicon Data. *Nature Methods* **13**, 581 (2016).
31 Lewis, Z. T. et al. Use of Bifidobacterial Specific Terminal Restriction Fragment Length Polymorphisms to Complement Next Generation Sequence Profiling of Infant Gut Communities. *Anaerobe* **19**, 62-69 (2013).
32 Ferry, B. *Breastfeeding* (International Statistical Institute
World Fertility Survey, 1981).
33 Greiner, T. Breastfeeding in Bangladesh: A Review of the Literature. *Bangladesh J Nutr* **10**, 37-50 (1997).
34 Akter, S. & Rahman, M. M. Duration of Breastfeeding and Its Correlates in Bangladesh. *Journal of Health, Population, and Nutrition* **28**, 595-601 (2010).
35 Whitehead, R. Nutritional Requirements: Infant Feeding Practices and the Development of Malnutrition in Rural Gambia. *Food and Nutrition Bulletin* **1**, 1-6 (1979).
36 (GBOS), T. G. B. o. S. (The Gambia Bureau of Statistics (GBOS). 2011).
37 Thorvaldsen, G. Was There a European Breastfeeding Pattern? *The History of the Family* **13**, 283-295 (2008).
38 Kersting, M. et al. [Breast Feeding Studies 1981-1983 in 1,500 Mothers in Dortmund and Haltern. Iii. Rates of Breast Feeding and Duration of Breast Feeding in the First Half Year]. *Monatsschrift Kinderheilkunde : Organ der Deutschen Gesellschaft fur Kinderheilkunde* **135**, 314-319 (1987).
39 Helsing, E. Supporting Breastfeeding: What Governments and Health Workers Can Do — European Experiences. *International Journal of Gynecology & Obstetrics* **31**, 69-76 (1990).
40 Verronen, P. Imetys on Muotia (Breast-Feeding Is Popular). *Suom Lääkäril* **39**, 1078-1079 (1984).
41 Brunn, S. in *World health forum 1986; 7 (1): 65-68.*
42 Tönz, O., Schwaninger, U., Holzherr, E. & Schafroth, M. [Infant Nutrition in Switzerland 1978. A Prospective Study on the Nutritional Habits During the First 6 Months of Life. I. Natural Nutrition: Breast Feeding]. *Schweiz Med Wochenschr* **110**, 937-947 (1980).
43 Bürger, B. et al. Breastfeeding Prevalence in Austria According to the Who Iycf Indicators—the Sukie-Study. *Nutrients* **13**, 2096 (2021).
44 Hasunen, K. Imeväisikäisten Ruokinta Suomessa Vuonna 2005. (2005).
45 Kohlhuber, M., Rebhan, B., Schwegler, U., Koletzko, B. & Fromme, H. Breastfeeding Rates and Duration in Germany: A Bavarian Cohort Study. *British Journal of Nutrition* **99**, 1127-1132 (2008).
46 Dratva, J., Gross, K., Späth, A. & Zemp, E. Swifs-Swiss Infant Feeding Study: A National Study on Infant Feeding and Health in the Child' S First Year: Executive Summary. *Basel: Swiss Tropical and Public Health Institute* (2014).
47 Wolf, J. H. Low Breastfeeding Rates and Public Health in the United States. *American Journal of Public Health* **93**, 2000-2010 (2003).
48 Ryan, A. S., Rush, D., Krieger, F. W. & Lewandowski, G. E. Recent Declines in Breast-Feeding in the United States, 1984 through 1989. *Pediatrics* **88**, 719-727 (1991).
49 Ryan, A. S. The Resurgence of Breastfeeding in the United States. *Pediatrics* **99**, e12-e12 (1997).
50 Callen, J. & Pinelli, J. Incidence and Duration of Breastfeeding for Term Infants in Canada, United States, Europe, and Australia: A Literature Review. *Birth* **31**, 285-292 (2004).
51 Rupnicki, S. Breastfeeding Report Card United States, 2020. (2020).
52 Health, U. D. o. & Services, H. Breastfeeding Report Card. *National Immunization Survey Available at: http://www cdc gov/breastfeeding/data/report_card2 htm Accessed April* **15** (2009).
53 Duranti, S. et al. Maternal Inheritance of Bifidobacterial Communities and Bifidophages in Infants through Vertical Transmission. *Microbiome* **5**, 66 (2017).
54 O'Leary, J. *Do Old Order Mennonites Hold the Key to Understanding Food Allergies*, <https://www.urmc.rochester.edu/news/story/do-old-order-mennonites-hold-the-key-to-understanding-food-allergies> (2017).

 doi:10.20944/preprints202202.0350.v1

55 Tannock, G. W. et al. Comparison of the Compositions of the Stool Microbiotas of Infants Fed Goat Milk Formula, Cow Milk-Based Formula, or Breast Milk. *Appl Environ Microbiol* **79**, 3040-3048 (2013).

56 Thompson, A. L., Monteagudo-Mera, A., Cadenas, M. B., Lampl, M. L. & Azcarate-Peril, M. A. Milk- and Solid-Feeding Practices and Daycare Attendance Are Associated with Differences in Bacterial Diversity, Predominant Communities, and Metabolic and Immune Function of the Infant Gut Microbiome. *Frontiers in Cellular and Infection Microbiology* **5** (2015).

57 Hassan, M. Z. et al. Antibiotic Use for Acute Respiratory Infections among under-5 Children in Bangladesh: A Population-Based Survey. *BMJ Global Health* **6**, e004010 (2021).

58 Akay, H. K. et al. The Relationship between Bifidobacteria and Allergic Asthma and/or Allergic Dermatitis: A Prospective Study of 0–3 Years-Old Children in Turkey. *Anaerobe* **28**, 98-103 (2014).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 February 2022 doi:10.20944/preprints202202.0350.v1

Figures and Tables

Table 1: Source of included cohorts, and number of infants included in the present study from each cohort

Country of Origin	**Published**	**Study**	**Enrollment based on intent to breastfeed?**	**Total Number Infants in Cohort**	**Number of Infants with Stool Samples**	**Number of Breastfed Infants at time of sample collection**
Austria	Partially	PASTURE	No	207	181	122
Bangladesh	Yes	Efficacy of Newborn Vitamin A Supplementation in Improving Immune Function (clinicaltrials.gov NTC01583972)	No	306	274	274
Finland	Partially	PASTURE	No	171	153	135
Gambia	Yes	Sub-study in The Early Nutrition and Immune Development (ENID) Trial, ISRCTN49285450	Yes	33	24	24
Germany	Partially	PASTURE	No	237	198	149
Switzerland	Partially	PASTURE	No	231	227	189
Davis, California, United States	Partially	UC Davis Lactation Cohort	Yes	95	60	60
Cincinnati, Ohio, United States	No	PREVAIL	No	245	45	26

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 February 2022 doi:10.20944/preprints202202.0350.v1

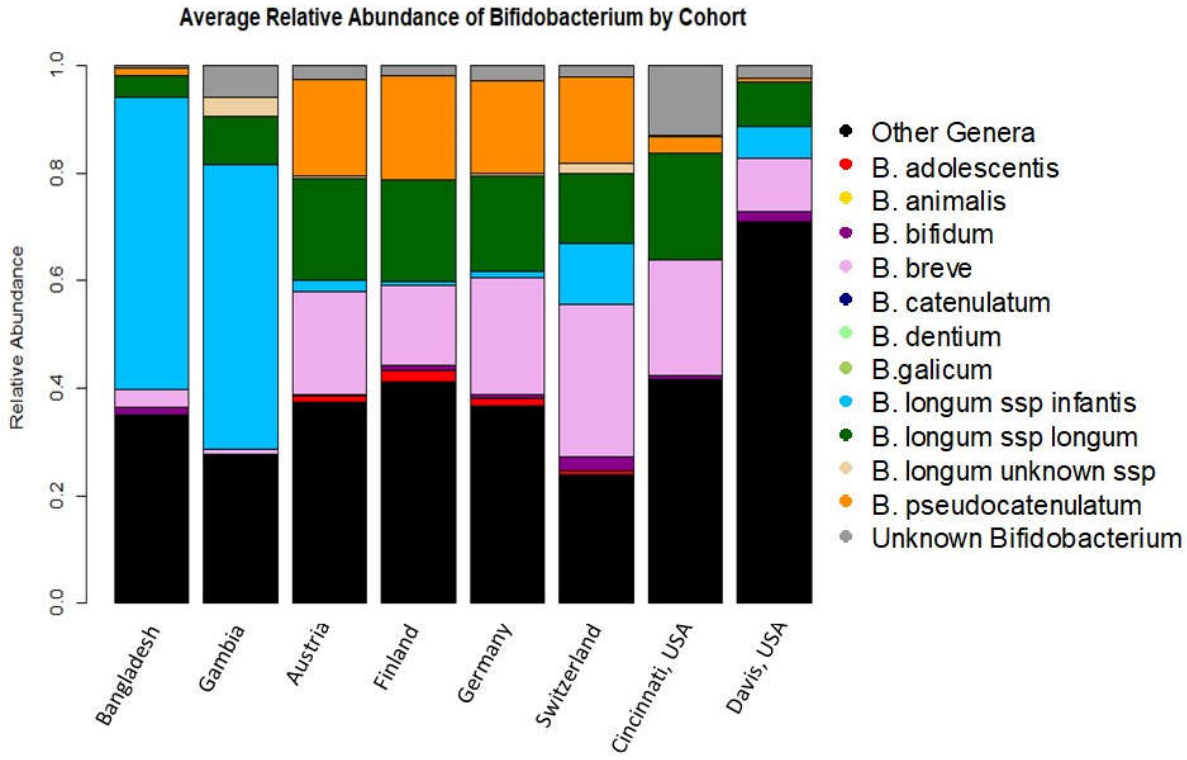

Figure 1: Average gut microbiome colonization patterns by country.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 February 2022 doi:10.20944/preprints202202.0350.v1

Table 2: Summary of *Bifidobacterium* prevalence in infants aged 1 to 2 months. Except for information on breastfeeding initiation and duration in a cohort, values are based solely on infants who were breastfed. The breastfeeding initiation (ever breastfed) rate and breastfeeding duration were calculated using data from all infants in a cohort with an available sample and a known breastfeeding status.

Cohort Location	Ever breastfed (median duration)	Historic breastfeeding pattern	Any Bifidobacterium	B. adolescentis	B. animalis	B. bifidum	B. breve	B. longum subsp infantis	B. longum subsp longum	B. longum subsp unknown	B. pseudocatenulatum
Bangladesh	100% (>2 years)	High	100%	5.8%	2.6%	32.5%	40.1 %	83.6%	26.3%	1.1%	8.4%
Gambia	100% (Unknown)	High	100%	0%	0%	0%	25%	91.7%	54.2%	12.5%	0%
Austria	91.8% (6.8 months)	Medium	100%	23.7%	0.82%	6.6%	55.7 %	4.1%	61.5%	4.9%	52.4%
Finland	99.3% (8 months)	Medium	100%	15.5%	0%	8.8%	40.0 %	0.74%	56.3%	3.7%	51.1%
Germany	91.1% (7.4 months)	Medium	100%	18.1%	0.67%	7.4%	61.1 %	4.0%	61.7%	4.0%	57.0%
Switzerland	97.3% (8 months)	Medium	100%	4.8%	0%	11.1%	58.7 %	14.8%	41.8%	6.3%	49.7%
Davis, California, United States	100% (9.3 months)	Low	65%	8.3%	0%	13.3%	36.7 %	8.3%	36.7%	1.7%	15%
Cincinnati, Ohio, United States	86.5% (3.1 months)	Low	97%	11.5%	0%	19.2%	61.5 %	0%	69.2%	0%	19.2%

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 February 2022 doi:10.20944/preprints202202.0350.v1

Table 3. Summary of GEE models comparing presence or absence of *Bifidobacterium* species in individual infants by cohort history of breastfeeding. Reference group was the high breastfeeding group. To account for multiple comparisons, a p-value less than 0.0062 is considered significant. The model for *B. animalis* failed to run, because after excluding infants who were not breastfed at time of sample collection, there were no infants from a low historical breastfeeding pattern cohort colonized by *B. animalis*.

Cohort Breast-feeding pattern.	*B. adolescentis* Odds ratio (95% CI, p-value)	*B. animalis* Odds ratio (95% CI, p-value)	*B. bifidum* Odds ratio (95% CI, p-value)	*B. breve* Odds ratio (95% CI, p-value)	*B. longum* subspecies *infantis* Odds ratio (95% CI, p-value)	*B. longum* subspecies *longum* Odds ratio (95% CI, p-value)	*B. longum* unknown subspecies Odds ratio (95% CI, p-value)	*B. pseudocatenulatum* Odds ratio (95% CI, p-value)
Medium	**4.1 (1.6 - 11, p=0.0041)**	NA	**0.22 (0.16 - 0.31, p<0.0001)**	**2.0 (1.3 - 3.1, p=0.0019)**	**0.010 (0.0037 - 0.029, p<0.0001)**	2.2 (0.98 - 5.0, p=0.055)	2.8 (0.92 - 8.2, p=0.069)	**12 (10 - 13, p<0.0001)**
Low	2.4 (1.0 - 5.6, p=0.041)	NA	**0.40 (0.28 - 0.58, p<0.0001)**	1.4 (0.70 - 2.9, p=0.32)	**0.0084 (0.0024 - 0.029, p<0.0001)**	1.8 (0.57 - 5.7, p=0.31)	0.63 (0.16 - 2.4, p=0.51)	**2.1 (1.8 - 2.5, p<0.0001)**

Table 4. Table of number of European infants colonized at any level by *B. longum subspecies infantis* at age 1 year by breastfeeding status. Prevalence of *B. longum subspecies infantis* appears to decrease rapidly during the weaning period, supporting the assumption that vertical transmission will be rare. P-values are for the Chi-square tests comparing *B. longum* subsp. *infantis* detection in still breastfed vs. no longer breastfed infants.

Country	Breast Fed at 1 Year (European countries) or 2 Years (Bangladesh)	B. longum subspecies infantis detected	B. longum subspecies infantis NOT detected	Percentage of infants colonized
Austria (p=0.003)	Yes	5	14	36%
	No	2	70	2.8%
Finland (p=0.65)	Yes	1	37	2.6%
	No	0	91	0%
Germany (p=1)	Yes	1	18	5.3%
	No	2	89	2.2%

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 February 2022 doi:10.20944/preprints202202.0350.v1

Switzerland (p<0.0001)	Yes	10	24	29%
	No	6	135	4.3%
Bangladesh (p=0.35)	Yes	13	51	20.3%
	No	5	40	11.1%

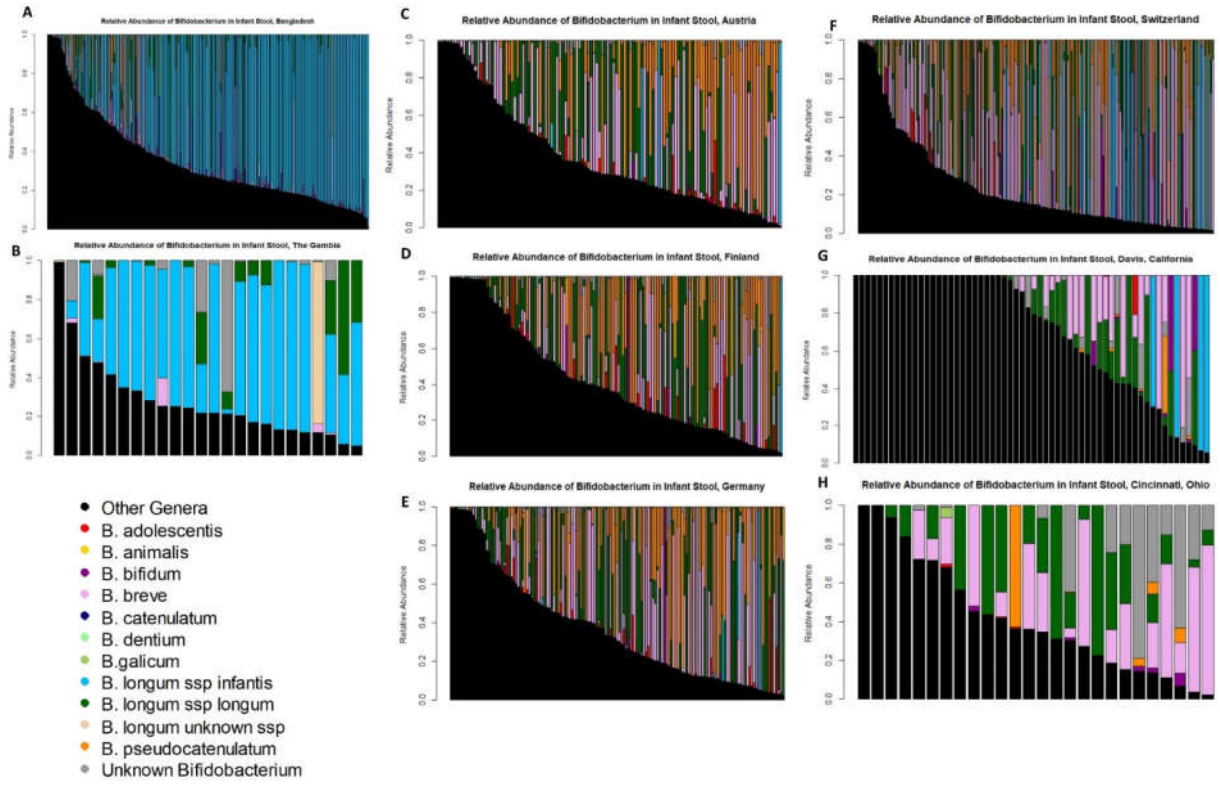

Figure 2. Relative abundance of different *Bifidobacterium* species in infant stool. *B. longum ssp infantis* is in blue, non-*Bifidobacterium* taxa are black. (A) Bangladesh (B) The Gambia (C) Austria (D) Finland (E) Germany (F) Switzerland (G) Davis, CA, USA (H) Cincinnati, OH, USA.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 February 2022 doi:10.20944/preprints202202.0350.v1

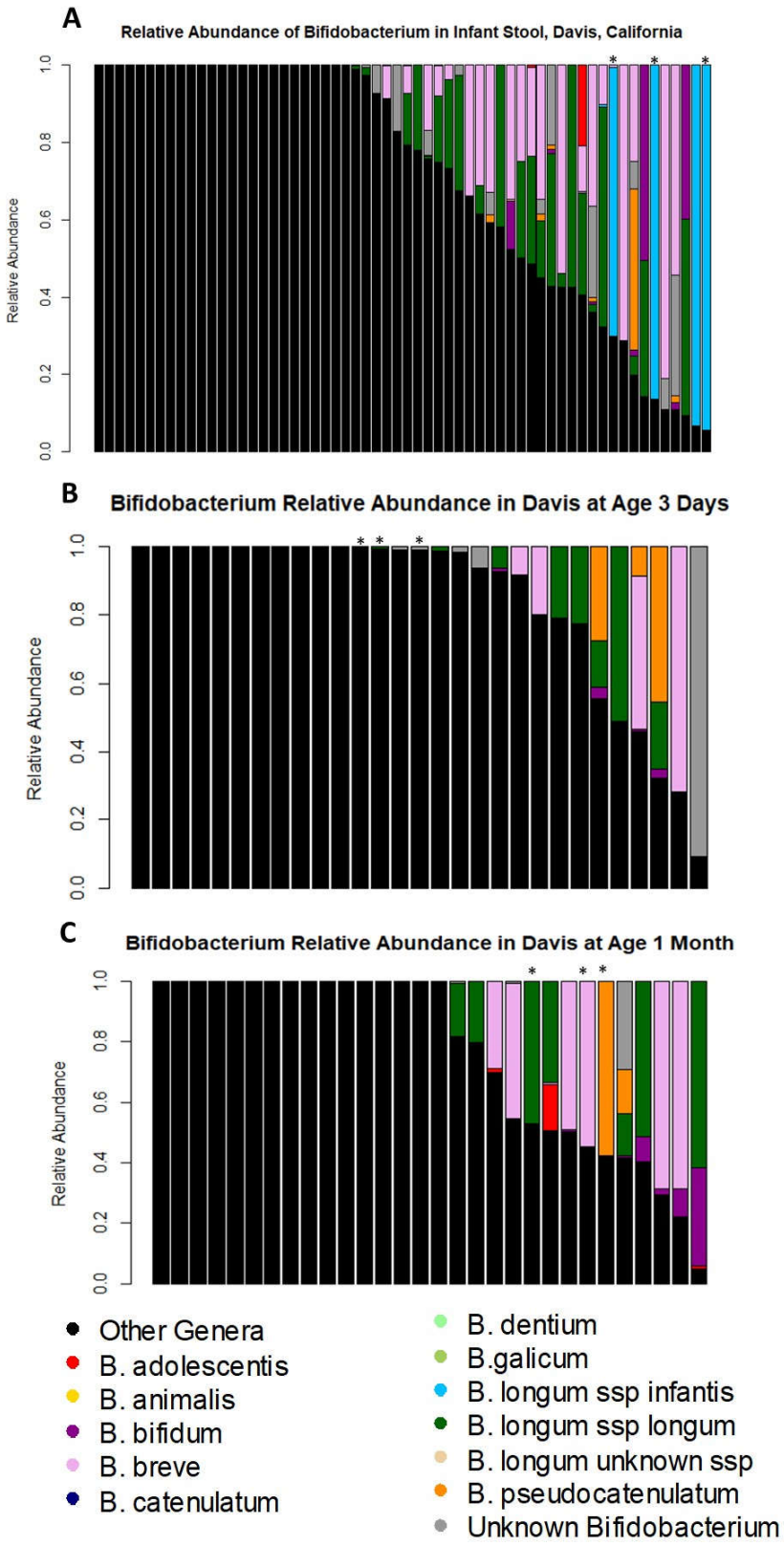

Figure 3. *Bifidobacterium* colonization in Davis, CA infants over time. Stars indicate samples from infants who were colonized by *B. longum* subsp. *infantis* at age 2 months. (A) Age 2 months samples (repeated from Figure 2) (B) Age 3 Days (C) Age 1 Month.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 February 2022 doi:10.20944/preprints202202.0350.v1

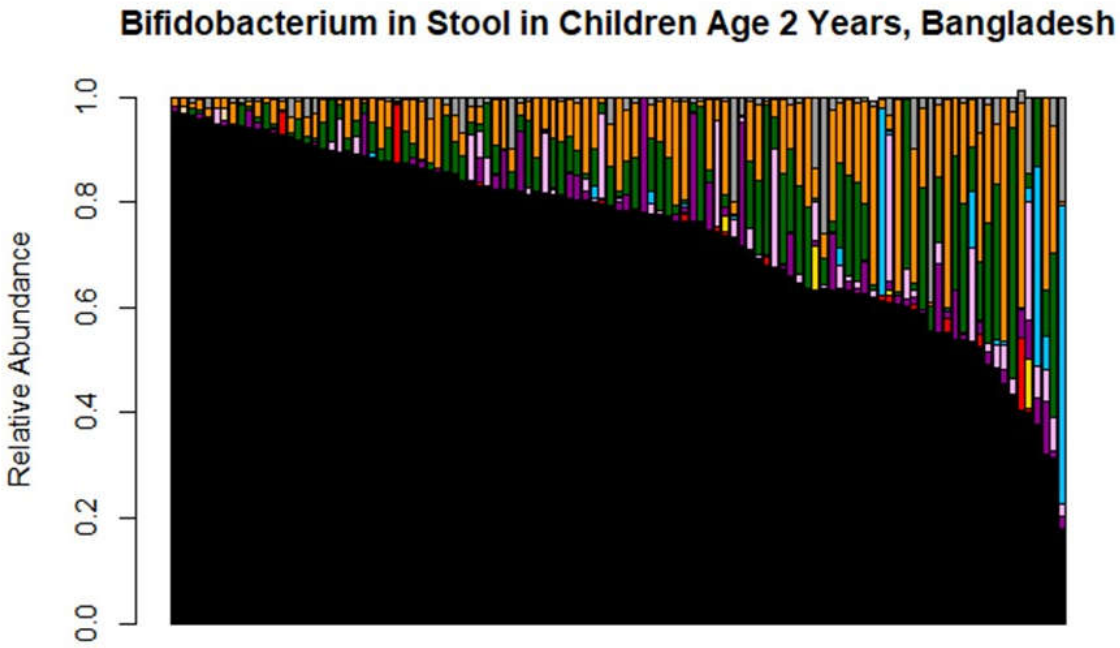

Figure 4. Bifidobacterium colonization in Bangladeshi cohort in infants age 2 years. Slight deviations from 100% are related to rounding errors in the Bif-TRFLP calculation of percentages. Note the reduced prevalence of *B. longum* subspecies *infantis* in blue; this is a subset of the same infants shown in Figure 2B.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 February 2022 doi:10.20944/preprints202202.0350.v1

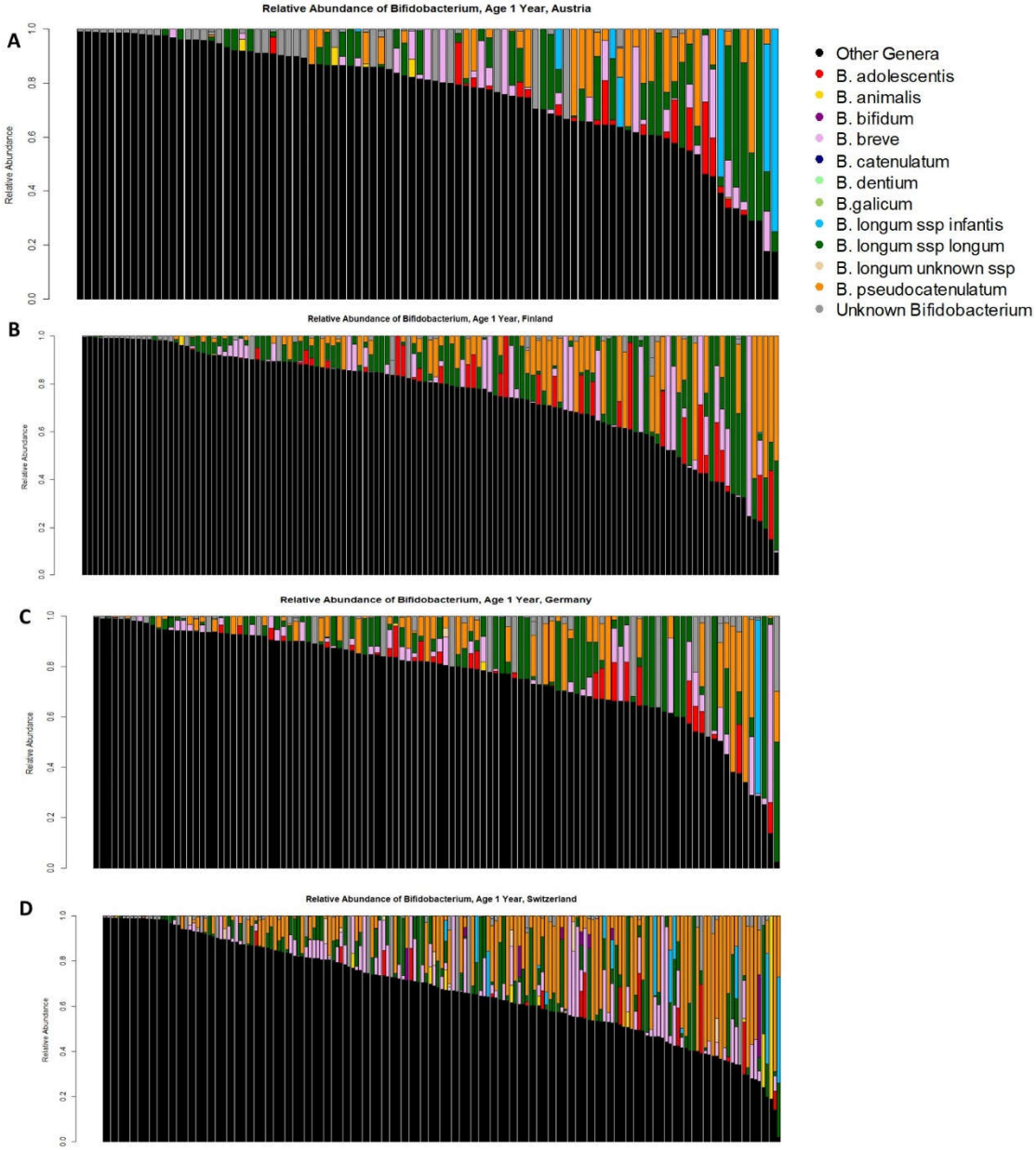

Figure 5. *Bifidobacterium* colonization in infants at age 1 year in (A) Austria (B) Finland (C) Germany and (D) Switzerland.