Probiotic supplementation accelerates gut microbiome maturation and reduces intestinal inflammation in extremely preterm infants

Jumana Samara[1,2,3,4], Shirin Moossavi[1,2,3,5], Belal Alshaikh[2,6], Van A. Ortega[1,2,3], Veronika Kuchařová Pettersen[1,2,3,7], Tahsin Ferdous[1,2,3], Suzie L. Hoops[8], Amuchou Soraisham[2,6], Joseph Vayalumkal[2], Deonne Dersch-Mills[2,6], Jeffrey S. Gerber[9,11], Sagori Mukhopadhyay[10,11], Karen Puopolo[10,11], Thomas A. Tompkins[12], Dan Knights[8], Jens Walter[13], Harish Amin[2,6]*, Marie-Claire Arrieta[1,2,3]*

[1] Department of Physiology and Pharmacology, Snyder Institute for Chronic Diseases, University of Calgary, AB. Canada
[2]Department of Pediatrics, Alberta Children's Hospital Research Institute, University of Calgary, AB. Canada
[3]International Microbiome Centre, University of Calgary, AB. Canada.
[4]Health Sciences Centre, Winnipeg, MB. Canada.
[5]Microbiome and Microbial Ecology Interest Group (MMEIG), Universal Scientific Education and Research Network (USERN), Calgary, Canada
[6]Calgary Zone Section of Neonatology, Calgary, AB. Canada
[7]Department of Medical Biology, UiT The Arctic University of Norway, Tromsø, Norway
[8]Biotechnology Institute and the Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN. USA
[9]Division of Infectious Diseases at Children's Hospital of Philadelphia, Philadelphia, PA. USA
[10]Newborn Care at Children's Hospital of Philadelphia, Philadelphia, PA. USA
[11]Department of Pediatrics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA. USA
[12]Lallemand Health Solutions, Montreal, QC. Canada
[13]School of Microbiology, Department of Medicine, and APC Microbiome Ireland. University College Cork, Ireland

*equal senior authorship

Correspondence:

marie.arrieta@ucalgary.ca

Summary

Probiotics are increasingly administered to premature infants to prevent necrotizing enterocolitis and neonatal sepsis. However, their effects on gut microbiome assembly and immunity are poorly understood. Using a randomized intervention trial in extremely premature infants, we tested the effects of a probiotic product containing four strains of *Bifidobacterium* species autochthonous to the infant gut and one *Lacticaseibacillus* strain on the compositional and functional trajectory of microbiome. Probiotic treatment accelerated the transition to a mature, term-like microbiome with higher stability and species interconnectivity. Besides infant age, probiotic *Bifidobacterium* strains and stool metabolites were the best predictors of microbiome maturation, and structural equation modeling confirmed probiotics as a major determinant for the trajectory of microbiome assembly. Probiotic-driven microbiome maturation was also linked to an improved, anti-inflammatory intestinal immune milieu. This demonstrates that *Bifidobacterium* strains function as ecosystem engineers that lead to an acceleration of microbiome maturation and immunological benefits in extremely premature infants.

Keywords: Premature infant, microbiome, mycobiome, probiotic, maturation, immune priming

Introduction

Postnatal microbial colonization in humans results in a dynamic assembly process that establishes the gut microbiota in a series of ecological succession events[1-3]. In infants born by vaginal delivery at term, early predominance of facultative anaerobic bacteria (i.e. *Streptococcus* spp., Enterobacteriaceae, *Staphylococcus* spp.) is followed by a community dominated by *Bacteroides* and *Bifidobacterium* species that further diversifies during and after weaning[2,4]. This process is drastically altered in infants born prematurely, with the magnitude of alterations correlating with the severity of prematurity[5-12]. Premature infants display a gut microbiome of reduced alpha-diversity, delayed colonization with obligate anaerobic bacteria and increased abundance in potentially pathogenic bacteria[5-12]. Despite a large degree of temporal and interindividual variability, the gut microbiome of the premature newborn follows patterns of microbial colonization that are to some degree conserved[3,6,7]. For example, extremely premature infants between 24-28 weeks gestational age (GA) are initially colonized by a community dominated by *Staphylococcus* spp., followed by *Enterococcus* spp. predominance between 28-32 weeks GA. Members of Enterobacteriaceae bloom later through interactions with *Staphylococcus* spp. between 32-35 weeks GA[3]. Following this period of facultative anaerobes predominance, strict anaerobic *Bifidobacterium* species become highly abundant at the age of term, when the premature microbiome begins to resemble the term infant composition[6,7].

The ecological drivers that disrupt the gut microbiota in premature infants are insufficiently understood. It has been proposed that organ-specific immaturity of preterm infants might provide selective pressure different from that of the term infant, either selecting for specific organisms and/or constitute habitat filters that prevent the colonization of the normal pioneer colonizers of the term infant gut[13]. Additionally, preterm infants are more likely to be

born by Caesarean section (C-section), receive antimicrobial treatment, achieve enteral feeding more slowly and require longer hospitalization compared to those born at term, all of which constitute potential determinants of microbiome alterations[14]. The consequences of the delayed microbiome maturation are also not well understood. Microbiome development in preterm babies is strongly correlated with GA, and the maturational delays may therefore reflect adaptations of the microbiota that are specific and perhaps necessary for preterm babies. However, extremely premature infants are strongly predisposed to devastating conditions like necrotising enterocolitis (NEC) and neonatal sepsis[15-17], which are not only linked to an altered gut microbiome[5,16] but can further be prevented through probiotics[18,19]. Given that probiotics modulate the microbiome in premature infants[20,21], their established benefits support a causal role for microbiome alterations as a true dysbiosis[22] in the etiology of these pathologies.

Probiotics are increasingly administered in neonatal intensive care units (NICUs) given their clinical effectiveness in reducing the risk of NEC and sepsis[18,19]. However, their use remains a matter of debate[23,24], and very little is known on the effect of probiotics on the assembly process of this nascent ecosystem and infant immune status. A recent study in term infants demonstrated that *B. infantis* EVC001 stably engrafts and dominates the community[25], and supplementation induced anti-inflammatory effects in term, breastfed infants[26]. However, it is unclear if probiotics exert the same effects in extremely premature infants who present with a much higher degree of dysbiosis and are at a heightened risk of infection and acute inflammatory conditions[15-17]. In addition, healthy infants are often colonized by a mix of *Bifidobacterium* species (*B. breve*, *B. bifidum*, *B. longum*) that can establish trophic interactions between themselves[27] and other genera[28] which might constitute the basis for robust community assemblies early in life[29].

Here we report findings from a randomized clinical trial of 57 extremely premature infants born at less than 1000 grams birth weight and less than 29 weeks GA (ClinicalTrials.gov Identifier: NCT03422562). Twenty-six infants were randomized to a probiotic treatment (FloraBABY, Renew Life®, Canada) containing four *Bifidobacterium* strains from species that are common and dominant in the infant gut [*B. breve* HA-129, *B. bifidum* HA-132, *B. longum* subsp. *infantis* HA-116 (*B. infantis* HA-116) and *B. longum* subsp. *longum* HA-135(*B. longum* HA-135)], and *Lacticaseibacillus rhamnosus* HA-111, and 31 infants were left untreated. Before, during, and 6 months after the intervention, we determined the presence and persistence of the probiotics using strain-specific qPCR, evaluated the bacterial and fungal microbiome using 16S and ITS rRNA sequencing and metabolomics, and measured cytokine levels in stool. We integrated these data through ecological and statistical models to determine the consequences of probiotic use on premature microbiome assembly and intestinal immunity.

Results

Bifidobacterium strains but not L. rhamnosus can stably colonize the premature infant gut.

Extremely premature NICU-resident infants were randomized to receive daily administration of FloraBABY or no probiotic. Probiotic administration started during the first week after birth following the collection of the first stool sample (T1), while two fecal samples were collected during treatment (T2 and T3), followed by a 2-week washout phase at term age (T4). A final sample was collected at 6 months corrected age (CA; T5) (**Figure 1A**). Two infants received probiotics prior to sample collection and thus their T1 samples were removed from the analysis.

Strain-specific qPCR showed increased fecal cell numbers for all strains during probiotic administration at timepoints T2 (2-3 weeks of age) and T3 (4-5 weeks of age) when compared to

the control group (**Figure 1B-F, Extended Data Figure S2)**. All probiotic strains remained significantly higher in the treatment group at T4 (2 weeks after administration). At T5 (6 months CA), all *Bifidobacterium* strains except *B. infantis* HA-116 remained significantly elevated in the treatment group (**Figure 1B-F; Extended Data Table S1**). While several infants still harboured detectable levels of *B. infantis* HA-116 at T5, cell numbers of *L. rhamnosus* HA-111 dropped below detection levels at T5 in all infants. These findings indicate stable colonization and proliferation of all *Bifidobacterium* strains in the premature infant gut for 6 months after administration was stopped, while *L. rhamnosus* HA-111 was unable to engraft (**Figure 1B**). Interestingly, *B. bifidum* HA-132, *B. longum* HA-135 and *B. breve* HA-129, but not *B. infantis* HA-116 or *L. rhamnosus* HA-111, increased to detectable levels in 93%, 53%, and 71% of control infants by 6 months CA, respectively (T5; **Figure 1B-F; Extended Data Table S2**), suggesting that transfer of these three probiotic strains to some control infants did occur during later stages of hospitalization.

Probiotics accelerate microbiome maturation in extremely premature infants to a level comparable with term infants

Previous observational studies have shown that probiotics can be used to modify the premature infant microbiome, mainly increasing alpha-diversity and the relative abundance of *Bifidobacterium* species[20,21]. However, the ecological effects on gut microbiome assembly and successional trajectory have not been systematically determined using an intervention trial. To achieve this, we applied an unsupervised clustering approach to the microbiome data collected temporally throughout the study. This analysis revealed four microbiome community types (C1-C4) (**Figure 2A and Extended Data Figure S3A**). Community type C1 and C2

dominated at T1, while C4 is completely absent at T1 but dominated at T5 (**Figure 2B**). There was a gradual increase in alpha diversity (Chao1) and community homogeneity as the microbiota matured from C1 to C4 (**Figure 2C and Extended Data Figure S3D**). Furthermore, C4 community type is characterized by high levels of *Bifidobacterium* while the less mature community types are dominated by *Staphylococcus* and *Enterobacteriaceae* (**Extended Data Figure 4C**), reflecting preceding succession stages in microbiome development[1-3].

To determine to what degree the community types detected in preterm infants differ to the microbiome of term infants, we compared them to microbiomes from 1-week (N=44) and 6-months (N=24), breastfed infants born at term. Ordination analysis based on Bray-Curtis dissimilarity showed that while the overall composition of the premature microbiome differed from term infants (**Extended Data Figure 4A**), microbiomes from community type C4 showed substantial overlap (on PCoA1) with the microbiome of term born infants (**Figures 2D and Extended Data 4B**). These findings establish that the community types detected in premature infants represent gradual stages of maturation of the gut microbiota that range from an immature microbiome to one that more closely resembles that of term infants.

An analysis of the impact of probiotics on community maturation revealed that there was no difference in community type distribution between the probiotics and control groups before treatment started during the 1st week of life (T1), with both groups consisting of C1 and C2 in equal proportions (**Figure 2B**). During the treatment period, which spanned from 2-6 weeks of age (T2-T3), community type C1 transitioned to C2 or C3 in both groups, but there was a proportion of infants only in the probiotics group that transitioned to C4 (**Figure 2B**). Infants in both control and probiotic groups predominantly consisted of C4 community type at 6 months CA (T5; **Figure 2B**). While the control group exhibited a delayed maturational pattern of gut

microbiome similar to what has been previously described in premature infants[3,6,7], 36% of the infants who received probiotics, arrived at the mature C4 community as early as T2 compared to none of the controls (**Extended Data Figure S3B-C**). This acceleration in microbiome maturation through the probiotic treatment was also seen in the Bray-Curtis analysis, where the average dissimilarity to full term microbiomes was lower at time points T2 ($p < 0.001$), T3 ($p < 0.001$), and T4 ($p = 0.014$) when compared to term breastfed infants, demonstrating restoration of the community (**Figure 2E-F**).

Probiotics promote a community with higher species interconnectivity and stability

Primary succession patterns in macro- and microbial ecology often follow an increase in community diversity and interaction network complexity[30,31]. In accordance, we observed increased species richness (**Figure 2C**). To further assess community ecological parameters, we determined interconnectedness, complexity, stability, and probabilities of transition between community types.

Network analysis revealed that inter-connectivity increased from C1 to C4 (**Figure 3A and 3C**). This ecological shift is strongly influenced by the probiotic intervention with a higher community interconnectivity in the treatment group as compared to the untreated controls (**Figure 3B**). Markov chain analysis to determine the probability of transitions between community types revealed that both the probability of the community to mature to C4, as well as to remain as C4, was higher in the probiotic group, indicative of higher community stability (**Figure 3D**). A time-to-event analysis confirmed that infants who were supplemented with probiotics showed a higher probability to mature to C4 earlier than controls, and that these effects persist beyond cessation of the probiotic (**Figure 3E**). Finally, a multivariate logistic

regression analysis showed that the impact of probiotics on the acceleration of microbiome maturation was more prominent than that of infant age, and other factors identified as microbiome-modulating factors in early life, including birth mode, feeding, and antibiotics[14] (**Figure 3F**). Together, this analysis indicates that probiotic supplementation to premature infants accelerates microbiota assembly towards a more mature and stable microbiome.

Probiotics accelerates gut metabolome maturation in extremely premature infants

We carried out untargeted metabolomics on a subset of fecal samples (N=82) to compare the intestinal metabolic milieu between infants who received probiotics and controls. Using permutational multivariate analysis of variance (PERMANOVA) on Bray-Curtis dissimilarities among samples, we identified that infant age and probiotics had strong effects on the premature infant metabolome composition, with sampling timepoint and probiotic intervention explaining 26.3% and 6.7% of the metabolome variance, respectively. ($P<0.001$; **Figures 4A-B)**. We also identified differences in temporal metabolic transition influenced by probiotic intervention and confirmed an interaction effect between timepoint and probiotic use on the metabolome (PERMANOVA, $R^2=8.4\%$, $P=0.03$; **Figures 4A**). We noted a transition in the metabolome as timepoints increased, and this transition was accelerated in infants who received probiotics (**Figure 4B**). All but T1 samples clustered together in the probiotic group, in contrast to control samples, in which the transitions were more temporally distinct. This suggests that this probiotic intervention not only accelerated the transition to a more mature microbiome composition, but also resulted in a more mature metabolic state.

To determine the metabolic characteristics of a mature microbiome in preterm infants, we compared the fecal metabolome of C4 (N=25) infants with that of the immature states (C1-C3,

N=27). Microbiome maturation (C4) made a significant contribution to variation in metabolome composition (R^2=7.3%, P<0.001; **Figure 4C**). Out of the 82 metabolites measured, we identified 14 differential metabolites as significantly different (Fold change >2, FDR P<0.05) **(Figure 4D and Extended Data Table S5).** These included elevated levels of the essential amino acids leucine, valine and phenylalanine, and the fatty acids oleic acid, palmitoleic acid and arachidic acid, in samples categorized as immature, suggesting the presence of nutritional substrates that remain unutilized by the immature microbiome and/or the premature gut.

We also compared the metabolic profiles of the immature and mature microbiome in preterm infants to those of infants born at term (N=30). Among the 14 metabolic features that differentiated mature and immature community states in preterm infants, 8 metabolites in the mature microbiome preterm group reached similar levels to term infants (**Figure 4E**). These included an increase in cholate and taurine in the mature microbiome composition. Cholate is a primary bile acid produced in high concentrations in the liver, and when conjugated with taurine forms taurocholic acid, the highest concentrated bile acid in bile[32]. Critical for fat digestion and absorption, bile acids are typically reduced in serum and duodenal aspirates in premature infants and they increase with postnatal age[33]. A mature microbiome composition also resulted in reduced levels of oleic acid (**Figure 4E**), the fatty acid found in highest concentration in breast milk[34], suggesting improved fat absorption, potentially from increased bile acid production in premature infants with a mature microbiome composition. We also detected a decrease in 3-nitrotyrosine linked to the mature microbiome composition, which approximated levels detected in term infants (**Figure 4E**). This metabolite is an established marker of cell damage, inflammation and nitric oxide production and it is elevated in a large number of pathological inflammatory diseases[35], including prematurity-related pathologies such as pulmonary

dysplasia[36,37] and NEC[38], further supporting the benefits of microbiome maturation in extremely premature infants.

L-cysteine, an important substrate for bifidobacteria (which are auxotroph for it[39]), was reduced in the mature microbiomes (**Figure 4E**), which may reflect L-cysteine consumption by microbial communities with a greater *Bifidobacterium* abundance. We also detected elevated levels of guanine, n-acetyl-DL-glutamic acid, and reduced creatine linked to microbiome maturity and reaching comparable levels to those in term infants (**Figure 4E**), which may also be the result of bifidobacteria. An increase in guanine and n-acetyl-DL-glutamic acid and a decrease in creatine were found in the stool of breastfed term infants compared to those fed formula[40,41], which correlated with the abundance of bifidobacteria[41]. These findings provide evidence for increased functional similarity between the mature preterm microbiome to that of the term breastfed babies, which are not explained by differences in breastmilk intake, as they were identical in the probiotic and control groups (**Extended Data Tables S3 and S4**). Finally, when comparing metabolite levels using features with the largest differences according to maturation state (highest fold-change values), the mature preterm samples more closely approximated the term metabolome than the immature preterm samples (**Figure 5F**). Altogether, these findings indicate that microbiome maturation in preterm infants results in potentially beneficial metabolic changes with important similarities to the intestinal metabolic milieu of healthy, breastfed infants born at term.

Bifidobacterial probiotic strains and metabolites drive microbiome maturation

To determine the drivers of microbiome maturation, we applied a random forest classifier to identify variables that can predict maturation to community type C4 (versus C1-C3), and their

relative importance. We included variables known to be major drivers in microbiome assembly[14], such as host (age, GA, sex), clinical (peri- and postnatal antibiotics, birth mode), dietary (breast(milk) feeding, hydrolyzed protein formula, fortification), as well as microbiome variables (probiotic strains cell numbers, probiotic duration) and differential fecal metabolites as variables for predictions. Apart from the infants chronological age, which was the best predictor, levels of creatine, taurine, guanine, n-acetyl-DL-glutamic acid, and cell numbers of the probiotic *Bifidobacterium* strains constituted the most important factors predicting gut microbiome maturation status (**Figure 5A),** showing higher Gini indices than factors often considered important, such as antibiotic treatment, birth mode, breast feeding, and GA. The *L. rhamnosus* HA-111strain grouped lower than these factors, further suggesting a lower effect of this strain in microbiome maturation in this clinical trial.

We also used structural equation modeling (SEM) to incorporate a theoretical framework of causal pathways underlying the associations between study variables and the premature gut microbiome (**Figure 5B**). Only time points T1-T4 were included in the model due to the reduced number of samples collected at T5 and the necessity to include complete sample numbers at each time point for SEM. We selected variables with a reported effect on the infant microbiome[14], including birth mode, GA at birth, antibiotic use, breast milk intake and probiotic use. Given the widespread use of breast milk instead of formula at the NICU where the study took place, breast milk intake could only be evaluated at T2, at a time when some of the infants received formula.

SEM analysis revealed that C-section and GA at birth were directly associated with bacterial richness at T1 ($\beta = -0.48$; $p<0.001$ and $\beta = -0.28$; $p=0.04$, respectively). Breast milk intake was directly associated with T2 richness ($\beta = 0.17$; $p=0.006$), yet a more prominent effect was observed for probiotics at T2 ($\beta = 0.595$, $p <0.001$). Although probiotics were being

administered at both T2 and T3, the effect on microbiome richness (Chao1) was not significant at the T3, yet microbiome composition at T2 strongly impacted subsequent communities' richness at T3 and T4 ($\beta = 0.74$; $p<0.001$ and $\beta = 0.62$; $p<0.001$, respectively; **Figure 5B**). This intriguing observation suggests that by impacting microbiome composition at an early time point (T2), probiotics may contribute to the trajectory of microbiome assembly, possibly through priority effects[42]. Similar significant effects were also made for alpha-diversity (Shannon index; not shown). Overall, these findings, together with the facts that probiotics persisted long after consumption ceased (**Figure 1B**) and that duration was not a strong predictor of microbiome maturation in the random forest model (**Figure 5A**), challenge the requirement of long-term probiotic administration to achieve compositional changes in the microbiome of extreme premature infants.

Probiotic use depletes Candida spp. but probiotic-Candida interactions do not modulate microbiome maturation

Given that multi-kingdom microbe-to-microbe interactions have been identified as drivers of the assembly process[3], we studied the temporal changes of the premature mycobiome and its association with probiotic use. Compared to what has been established for the bacterial microbiome[6,7], temporal analysis of the premature gut mycobiome did not reveal major shifts in the relative abundance of the most abundant fungal genera between T1-T4 (**Figure 6A**). Community typing also identified four fungal clusters yet these did follow distinct patterns of community transition (**Extended Data Figure S5A-B**), suggesting that the gut mycobiome may not display community maturation patterns in the same manner as bacterial communities. Probiotic administration resulted in a significant decrease in the relative abundance of *Candida*

spp. (**Figure 6A-B**), in agreement with previous studies[43,44]. While many samples had low relative abundance of *Candida* spp. in our study, more samples were dominated by very high levels of *Candida* spp. in the infants who did not receive probiotics (**Figure 6C**). When categorising at a 50% relative abundance threshold, the proportion of samples from infants with >50% *Candida* spp. abundance was significantly lower in the intervention group (**Figure 6D**), indicating that probiotic use induces a strong anti-*Candida* effect.

We assessed the specific role of *Candida* spp. as a modulator of the effect of probiotic use on gut microbiome maturation. We used SEM to evaluate the direct influence of *Candida* spp. abundance on bacterial richness (**Figure 6E**), as well as its indirect role on microbiome maturation via interactions with probiotic strains (**Figure 6F**). While probiotics and milk type were significantly associated with the gut microbiome richness, we did not observe a direct association of *Candida* spp. with bacterial richness in this model (**Figure 6E**). Similarly, the association of probiotic strains with bacterial community types was not influenced by the relative abundance of *Candida* spp. (**Figure 6F**), denoting the stronger ecological influence of the probiotic strains compared to endogenous *Candida* sp. The strong anti-*Candida* effect of the probiotics may explain why this fungal species is not associated with the successional patterns observed in our study, as it was in a recent thorough ecological analysis of the premature microbiome assembly without a probiotic intervention[3]. Although the effect of the probiotic on *Candida* spp does not seem to constitute a mechanism by which microbiota maturation is enhanced, the effect is nevertheless important given the clinical relevance of *Candida* spp in nosocomial infections among premature infants[43,44].

Probiotic-induced microbiome maturation reduced proinflammatory cytokines in stool of extremely premature infants

Extremely premature infants are at an increased risk of NEC, a devastating inflammatory condition [15-17]. To investigate the effect of probiotics on intestinal inflammation, we determined the concentration of 17 cytokines and calprotectin in stool in a subset of samples (N=170). Cytokines play a central role in immune and inflammatory functions in the gut and are known to accumulate in stool and reflect intestinal inflammatory processes[45]. We applied generalized estimation equation models on longitudinal data to determine differences in stool cytokines during the time of hospitalization and after the probiotic intervention started (T2-T4). Probiotics led to an overall reduction in several important proinflammatory cytokines, including calprotectin, IFN-γ, IL-12p70, IL-4, as well as an increase in IL-22 (**Figure 7A**). In the gut, IL-22 exerts generally protective functions, such as maintaining barrier function and tissue injury regeneration[46], with recently reported critical role in the prevention and treatment of NEC in mice[47]. This demonstrates a strong and consistent intestinal anti-inflammatory effect of probiotics in extremely premature infants. (**Figure 7A and Extended Data Table S6**).

We also compared cytokine levels in relation to microbiome maturation (C4 vs. C1-C3 vs. Term). There was a significant decrease in IFNγ, IL-1β and IL-8 and calprotectin in stool samples from preterm infants with a mature microbiome composition compared to those with an immature microbiome composition, and the levels of IFNγ and IL-1βa in the mature microbiome were similar to those detected in term infant stool samples (**Extended Data Figure S6**). Overall, the differences between the immune status of preterm and term infants were significantly smaller for infants harboring the mature microbiome type (C4) as compared those with more immature microbiomes (C1-3) (**Figure 7C**). Finally, correlation analysis between cytokine concentrations

and microbial abundances revealed numerous significant correlations. Pathobionts, specifically the genera *Staphylococcus* and *Streptococcus*, showed positive correlations, while cell numbers of the probiotic strains showed negative correlations with the majority of the immune factors measured (**Figure 7B**). These findings suggest a predominant role of the probiotic strains on the immune milieu detected in stool samples from extremely premature infants.

Discussion

Microbiome maturation is disrupted and delayed in preterm infants predisposing the infant to life-threatening pathologies[15-17]. Our work demonstrated that a probiotic formulation leads to the stable colonization of *Bifidobacterium* strains weeks before bifidobacteria become dominant members of the fecal microbiome in untreated pre-term infants[6,7]. -This is in line with what was recently reported by Alcon-Gener et al.[20] in an observational study, showing strong and persistent colonization by *B. bifidum* after supplementation to preterm infants born at <32 weeks GA[20]. Our study further revealed that probiotics expedited transition to a more mature bifidobacterial-high community state with enhanced stability and species interconnectivity, two key features of later stages of primary succession[48]. Cell numbers of the *Bifidobacterium* strains administered with the probiotic and stool metabolites were among the strongest predictors of maturation, providing a mechanistic link between probiotic administration and an acceleration of microbiome maturation to a state more closely resembling the vaginally born, breastfed infant microbiome, the current benchmark for a desired term infant microbiome[14].

Although our study was not powered to capture health outcomes in this population, we detected favourable metabolic and immune features associated with probiotic-induced microbiome maturation. This includes favourable changes to features previously associated with

NEC in premature infants, including oleate[49], proinflammatory cytokines[50,51], and 3-nitrotyrosine[38]. Overall, the mature microbiome composition resulted in marked metabolic and immune differences that approximated the term stool metabolome (**Figures 4D and 7B**), and are indicative of improved fatty acid absorption, breastmilk metabolism, and reduced inflammation. Our findings complement the recently reported immune silencing effect of the probiotic *B. infants* EVC001 on term infants[26], extending the evidence for *Bifidobacterium* strains as drivers of beneficial immune imprinting during early life. These findings, as well as the ecological attributes of the more mature and stable microbiome suggest a beneficial effect to extremely premature infants, especially considering the well-established role of bifidobacteria excluding pathogenic organisms or providing cues for the developing immune system[52]. Larger studies in premature infants are needed to confirm if the metabolic and immune benefits resulting from an accelerated microbiome maturation result in improved health outcomes in extremely premature infants. Given that conditions such as NEC are driven by inflammation, such knowledge has substantial clinical implications.

The pronounced effects of probiotic administration on microbiome maturation can be explained using an ecological framework. To establish in the gut, organisms must first overcome the habitat filters present and then possess traits to acquire the available resources to become competitive[53]. In contrast to many other probiotic products, the probiotic used in this study is composed of *Bifidobacterium* strains from autochthonous species that naturally dominate the early-life microbiota of infants[29,52]. Such strains, in contrast to *L. rhamnosus* HA-111, are highly adapted to the infant gut. These adaptation include the ability to utilize human milk oligosaccharides and sugar hexoses[52], to competitively exclude other microbes, including pathogens through short chain fatty acid production[54], to decrease the intestinal luminal pH[52],

and to sustain metabolic cross-feeding of other gut microbiome species[55,56]. Our random forest analysis revealed that all *Bifidobacterium* strains (but not *L. rhamnosus* HA-111) contributed to microbiome maturation, suggesting a contribution of the wider *Bifidobacterium* community to microbiome assembly. The strongest predictor among the bifidobacteria, *B. bifidum*, provides substrates (fucose and sialic acid) from the hydrolysis of mucus and HMOs to other microbiome members[27,57,58], while the weakest predictor, *B. infantis*, internalises substrates without sharing[29], supporting a contribution of cross-feeding in microbiome maturation. Our findings further point to the importance of priority effects in that an earlier arrival of the probiotic strains enhances both their own persistence and modifies the trajectory of the assembly process[42,59]. Given the rapid and sustained ecosystem transformation linked to the probiotic *Bifidobacterium* strains, we propose that bifidobacteria act as ecosystem engineers[14] in the premature microbiome, capable of building, transforming, and preserving the microbial habitat in the infant gut.

Apart from providing strong evidence for the ability to use probiotics to restore the microbiome in preterm infants, our findings provide important clues on the ecological factors that lead to the pronounced disruptions observed in preterm microbiomes[5-12]. Our findings show that autochthonous *Bifidobacterium* strains can effectively and stably colonize the preterm gut. In addition, our random-forest analysis and structural equation modeling showed that such strains and metabolites associated with their predominance in the community are more important determinants of microbiome maturation than the host, clinical, and dietary factors often considered to play important roles. This suggests that the premature microbiome is not primarily disrupted through treatments and feeding practices of a modern NICU, or the premature physiological or immunological state of the host, and that microbiome maturational delays are unlikely to reflect necessary adaptations of the microbiota to the premature conditions. Instead,

our findings point to the inability of the premature infant to acquire the necessary strains to initiate the assembly process.

Ecologically, the human gut microbiota can be viewed as a meta-community in which individuals are linked through dispersal, which constitutes a key ecological process that shapes microbiome assembly at local scales[60]. Our strain-specific quantification showed that some infants in the control group did acquire the probiotic strains (Figure 1), likely because they were housed in the same NICU, demonstrating the ability to acquire early colonizers through horizontal transmission. However, this only occurred in a smaller subset of infants, and most infants acquired strains them later in microbiome development. These findings demonstrate that dispersal occurs infrequently in an NICU, possibly due to hygienic barriers to prevent infections, as well as the clinical practices linked to preterm births that disrupt vertical transmission from the mother to the infant (c-sections, antibiotics, maternal separation, etc.), all of which can reduce exposure to pioneer organisms that colonize term infants. This dispersal barrier may also contribute to immune dysregulation resulting in increased intestinal inflammation, which is central to the pathogenesis of inflammatory and infectious pathologies in extremely premature infants. If probiotics contain the right microbes that have evolved as early colonizers in humans, they can essentially function as a mechanism to restore the dispersal process. In this context, probiotics fall within the framework of ecological restoration as an attempt to reach a desired community, or to avoid an undesirable one. The findings of this study show that such an approach has great potential for clinical applications with health benefits to very vulnerable infant population.

Acknowledgements

The authors would like to acknowledge study participants and their families for their support to research, as well as NICU nurses for their effort in sample collection and communication with the study group. The following collaborators supported clinical and laboratory aspects of this research: Rachel Sheinfeld, Erik van Tilburg Bernardes, Mackenzie Gutierrez, Kristen Kalbfleisch, Julia Gorospe, Amanda Piano, Dr. Neha Bansal, Marija Drikic and Ryan Groves. This work was supported by funds from the Cumming School of Medicine, the Alberta Children Hospital Research Institute, the Snyder Institute of Chronic Diseases, the Canadian Institutes for Health Research. The International Microbiome Center is supported by the Cumming School of Medicine, University of Calgary, Western Economic Diversification and Alberta Economic Development and Trade, Canada. S Moossavi is supported by CIHR and Killam Postdoctoral Fellowship. V.A.O is supported by a MITACS Elevate Fellowship. J.W. acknowledges support by Science Foundation Ireland (SFI) through an SFI Professorship (19/RP/6853) and a Centre award (APC/SFI/12/RC/2273_P2) to the APC Microbiome Ireland.

Author Contributions

H.A., B.A., A.S., J.V and D. D-M contributed to the design of the premature RCT study. J.S. and B.A. monitored the clinical trial. J.S. communicated with study participant families, conducted study interviews and compilated all clinical data. J.S., V.K.P and V.A.O prepared all the samples for sequencing and metabolomics analysis. J.S carried out qPCR analysis. S.Moossavi, J.S. and M-C.A analyzed the 16S and ITS2 sequences. S.Moossavi, J.S., V.A.O and V.K.P analyzed the metabolomics data. S.Moossavi and T.F. carried out statistical analysis. V.A.O performed immune and protein determination assays. S.Moossavi and M-C.A created the paper figures.

T.A.T. provided probiotic strains and with J.W, guidance on qPCR protocol. S.H., D.K., J.S.G, S.Mukhopadhyay, and K.P. provided the 16S sequences from the MAGIC Study. M-C.A, S.Moossavi, J.S and J.W. contributed to data interpretation and writing the first and subsequent drafts of the manuscript. All authors edited the manuscript and contributed extensively to the work presented here.

Declaration of interests

T.A.T. was the Research Director at Lallemand Health Solutions, the manufacturers of FloraBABY. The other authors declare no competing interests.

Figure Legends

Figure 1. Probiotic strains can stably colonize the extremely premature infant gut. A) Study design for the randomized controlled trial of probiotics in extremely preterm infants. In the treatment group, probiotic was started in the first week of life before sample collection (T1) and continued until 37-39 weeks gestational age (GA) weeks spanning T2 and T3. Additional samples were collected after cessation of probiotic at 39-40 weeks GA (T4) and 6 months corrected age (CA) (T5). **B-F)** Concentration of probiotic strains assessed by strain-specific qPCR demonstrates increased concentration of all probiotics strains immediately after starting probiotic at T2. *Lacticaseibacillus rhamnosus* decreased after cessation of probiotics (B) while the *Bifidobacterium* strains showed stable colonisation until 6 months CA. The dashed line denotes the limit of detection (10^3 bacterial cells/ml). P values are obtained from linear mixed models (LMM) and post estimation for linear combination of coefficients (see also Extended Data Table 2). LOD, limit of detection.

Figure 2. Probiotics accelerate gut microbiome maturation in extremely preterm infants. **A)** Four gut microbiome community types were identified using hierarchical clustering on Bray-Curtis dissimilarity matrix. Association of the community types with beta diversity was tested using PERMANOVA. **B)** Microbiome community type distribution across timepoints and probiotic use. Community types showed temporal distribution, with C1 and C2 more frequent in earlier and C4 in later timepoints. As a result, C4 is considered the mature community type, which appeared earlier in infants treated with probiotics. **C)** Comparison of bacterial richness (Chao1) in community types (See Extended **Figure 2D** for comparison of beta diversity). **D)** Comparison of the maturational patterns of the microbiome community types with term infants

at 1 week and 6 months of age. **E-F)** Comparison of the temporal development of preterm infant microbiome with term infants at 1 week and 6 months of age in controls **(E)** and probiotic-treated infants **(F)**. Centroid of each timepoint is denoted as the red circle and the distance to the centroid of each timepoint to the centroid of 6-month term infants are presented as labels. Trend analysis in panels C-F were conducted using trendyspliner in SplinectomeR package[70].

Figure 3. Probiotics promote a microbial community with higher interconnectivity and stability. A-B) Network analysis of the preterm infant microbiome along the microbiome maturation trajectory **(A)** and by intervention **(B)**. **C)** Comparison of network degree and strength across community types. **D)** Probability of transition between community types assessed by Markov Chain modelling compared in controls and probiotic group. **E)** Time-to-event analysis demonstrates that probiotics accelerates transition into the C4 mature community type. Kaplan-Meyer curve for the probability of not reaching the mature community type is shown. **F)** Multivariable logistic regression demonstrating the association of probiotic treatment with microbiome maturation independently of early life events. Adjusted Odds Ratio (OR) and 95% confidence interval (CI) are presented for all variables in the model.

Figure 4. Probiotic-induced microbiome maturation is reflected in the stool metabolome. A-B) Principal component analysis of gut metabolome in premature infants at different timepoints and by intervention. Interaction between the effects of timepoint and probiotics was tested using PERMANOVA. **C)** Principal component analysis of gut metabolome in premature infants with mature (C4) vs. immature (C1-C3) community types. Effect of maturational status on the variance of the metabolome was tested using PERMANOVA. **D)** Differentially enriched

metabolites in mature (C4) vs. immature (C1-C3) community types as assessed by volcano plot with fold change threshold of 2 and adjusted t-test threshold of 0.05. Pink circles represent features above this threshold. **E)** The most discriminatory metabolic features from immature (gray) or mature (turquoise) microbiome maturation status in premature infants compared to term, breastfed infants (purple). Comparisons were made by pairwise Wilcoxon test. **F)** Metabolite levels by microbiome maturity in relation to term breastfed infants. Mean fold difference in the mature-term vs. immature-term comparisons are shown.

Figure 5. Probiotic strains and stool metabolites are predictive and drivers of microbiome maturation. A) Predictors of mature microbiome community type (C4 vs. C1 C2 & C3 combined) ordered by their importance identified through random forest modelling using 10-fold cross-validation, 500 trees, and 1000 permutations. **B)** Structural equation modelling was used to differentiate the influence of probiotics on bacterial richness (Chao1) at each timepoint while taking into account the structure of association of other early life factors. Probiotic was administered during T2 and T3 timepoints. Model fit was assessed using p value, CFI, RMSEA, and SRMR. Abx, antibiotics; CFI, comparative fit index; C/S, Caesarean section; RMSEA, root mean square error of approximation; SRMR, standardized root mean residuals.

Figure 6. Probiotic use depletes *Candida* spp. but probiotic-*Candida* interactions do not modulate microbiome maturation. A) Mycobiome community structure at genus level compared in controls and infants who received probiotics. **B)** Longitudinal analysis of *Candida* spp. according to the intervention using splinectomeR reveals significantly lower abundance in the probiotic group. **C)** Distribution of *Candida* spp. by intervention confirms lower average

relative abundance in the probiotic group. **D)** Categorizing *Candida* spp. relative abundance into < 50% or >50% revealed the infants who received probiotic are less frequently dominated by high levels of *Candida* spp. **E-F)** Structural equation modelling to examine the direct effect of *Candida* spp. on bacterial richness **E)** and indirect effect on microbiome maturation via interaction with probiotic strains **F).** Model fit was assessed using p value, CFI, RMSEA, and SRMR. CFI, comparative fit index; C/S, Caesarean section; RMSEA, root mean square error of approximation; SRMR, standardized root mean residuals.

Figure 7. Probiotic-induced microbiome maturation reduced proinflammatory cytokines in stool of extremely premature infants. A) Cytokine concentrations in premature infants according to the intervention. Comparisons were made by generalized estimating equation (Extended Data Table S6). **B)** Correlation of fecal cytokine levels with the 12 most abundant bacterial genera (mean relative abundance > 1%), *Candida*, and probiotic strains log 10 transformed cell numbers. Statistical significance was assessed by adjusting for multiple comparison using Benjamini and Hochberg method. **C)** Cytokine levels by microbiome maturity in relation to term breastfed infants. Mean fold difference in the mature-term vs. immature-term comparisons are shown.

576

577

STAR Methods

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Marie-Claire Arrieta (marie.arrieta@ucalgary.ca).

Experimental model and participants details

Inclusion and exclusion of study participants

This study was part of a randomized, open-label, controlled trial in the NICU of the Foothills Medical Centre in Calgary (ClinicalTrials.gov Identifier: NCT03422562). FloraBABY (Renew Life®, Canada) probiotic was administered to infants in the intervention arm after randomization. Eligible participants were premature infants admitted to the NICU with birth weight < 1000 grams and born at less than 29 weeks GA. Eligible infants were identified within 24 hours of birth and parents were approached for informed consent. Once consent was obtained, infants were randomly assigned in blocks of 4 to receive either FloraBABY probiotics or no product. Randomization was conducted using a computer-generated table of random numbers. The study excluded infants with major congenital anomalies, hypoxic-ischemic injury and NEC or bowel perforation occurring within 72 hours of birth. Probiotic administration was started before 7 days of age and continued until 37 weeks post-menstrual age, at a dose of 0.5g per day in 1 ml of milk or colostrum as part of the feeding. Each dose contained 4×10^9 total colony forming unit (CFU) of four *Bifidobacterium* strains (*B. breve* 1.2×10^9 CFU, *B. bifidum* 8×10^8 CFU, *B. infantis* 6×10^8 CFU, *and B. longum* 6×10^8) together with *Lacticaseibacillus (*formerly *Lactobacillus*[61]*) rhamnosus* 1×10^9 CFU, mixed with maltodextrin and ascorbic acid. No probiotic or placebo was given to infants in the control group. Treatment for the intervention group started after obtaining informed parental consent and after the first

stool sample was obtained, except for two infants, who received the probiotic before the first stool sample. Probiotics were administered until the age of term (37-39 weeks post-menstrual age). Total probiotic treatment duration ranged between 45-87 days, depending on gestational age at birth (**Figure 1A**). This trial was conducted in accordance and compliance with all relevant ethical regulations by the Conjoint Health Research Ethics Board of the University of Calgary (approved protocol REB16-0542).

Maternal, infant and early-life factors

The following variables were collected throughout the study and incorporated in the analysis: GA at birth, chronological age, birth weight, sex, number of older siblings, mode of delivery, maternal antenatal administration of antibiotics, age in days at the start of enteral feeds and inclusion in the study, total duration of probiotics, duration of neonatal antibiotic use (type and duration), type of feeds during NICU and up to 6 months CA, including milk type, fortification, and type of fortification.

METHOD DETAILS

Sample Collection and Processing

Stool samples were collected at five time points: prior to first probiotic administration (T1); 2-3 weeks after first administration (T2); 4-5 weeks after first probiotic administration (T3); 2 weeks after probiotic discontinued (T4); and at 6 months CA (T5; **Figure 1A**). CA refers to the infant age if the pregnancy would have gone to term. Stool samples for the control group were collected at matched gestational and chronological age time points. Stool was collected directly from the infant's diaper by NICU nurses (T1-T4) or participant parents at home (T5). Samples were

placed at 4°C in the NICU or at home for a maximum of 12 hours, or at -20°C in a NICU or home freezer for up to 48 hours and were stored at -80°C upon arrival in the laboratory for subsequent processing.

DNA extraction

DNA was extracted from ~50 mg of stool. Samples were mechanically lysed using MO BIO dry bead tubes (MO BIO Laboratories, USA) and the FastPrep homogenizer (TissueLyser II, Qiagen, Hilden, Germany) before DNA extraction with the DNeasy PowerSoil Pro Kit according to the manufacturer's instructions (Qiagen, Canada). Following extraction, DNA concentration was measured in a NanoDrop spectrophotometer (Thermofisher, Canada) and subsequently used in qPCR and sequencing reactions.

Quantitative PCR

To specifically quantify FloraBABY strains in fecal samples, qPCR was performed on genomic DNA using specific primer sequences (**Extended Data Table 2**) and qPCR protocols previously validated to detect these probiotic strains in stool samples[62]. We carried out further validation of the specificity of the primers using individual strains in maltodextrin powder and a standard operating procedure, provided by Lallemand Health Solutions, Montreal, Canada. Each strain powder was spiked into stool samples negative for the probiotic strains. These samples were obtained from infants enrolled in a longitudinal birth cohort study in rural Mexico, with no history of exposure to probiotics. To determine the concentration of each strain, one gram of lyophilized powder of each probiotic strain was diluted in 99 ml phosphate buffered saline to obtain 10^{-2} solution. Flow cytometry counts provided the concentration (bacteria/ml) to calculate the total count of cells in 10^{-2} solution for each strain. A selected set of stool samples from

644 Mexican cohort were spiked with the exact volume required to reach a concentration of 10^9
645 bacteria/ml. Unspiked stool samples were used as negative controls. To validate the qPCR
646 methods, ten-fold dilutions (10^2 to 10^9) of DNA extracted from the spiked and unspiked aliquots
647 were used as templates in qPCR validation plates (triplicates for each dilution). Reactions were
648 run using StepOne™ Real-Time PCR System using the following protocol: 2 initial steps of
649 2 min each at 50 °C and 95 °C, followed by 40 cycles of 15 seconds at 95 °C, 30 seconds at 60
650 °C and 30 seconds at 72 °C. DNA concentrations were measured for all five probiotic strains
651 using serially diluted spiked DNA extracted from spiked stool samples as standards. Clinical
652 samples were run on duplicate using 4ng of extracted DNA as template. Cell numbers were
653 calculated as cell/ml based on the standard curve method. Cell number values obtained below the
654 detection limit (10^3 cells/ml for all probiotic strains) were substituted with limit of detection
655 divided by square root of 2 to account for variance in statistical tests and models.

656 *16S rRNA and ITS2 gene sequencing*

657 PCR was used to amplify the V4 region of the bacterial 16S rRNA gene and the ITS2 region of
658 the fungal ITS genetic marker from fecal DNA. This generated ready-to-pool dual-indexed
659 amplicon libraries as described previously[63]. 16S and ITS amplicon libraries were prepared at
660 Microbiome Insights (University of British Columbia, Vancouver, Canada). In-house extracted
661 DNA samples were sent to the facility and amplified using Phusion Hot Start II DNA
662 Polymerase (Thermo-Fisher). PCR products were purified, and DNA concentration normalized
663 using the high-throughput SequalPrep Normalization Plate Kit (Applied Biosystems, USA) and
664 quantified accurately with the KAPA qPCR Library Quantification kit (Roche, Canada).
665 Controls without template DNA and mock communities with known amounts of selected
666 bacteria and fungi were included in the PCR and downstream sequencing steps to control for

667 microbial contamination and verify bioinformatics analysis pipeline. Samples were sequenced in
668 two runs and biological controls were included in both runs to assess for batch effects. The
669 pooled and indexed libraries were denatured, diluted, and sequenced in paired-end modus on an
670 Illumina MiSeq (Illumina Inc., San Diego, USA). 16S rRNA and ITS2 gene sequencing were
671 performed at Microbiome Insights, Vancouver, BC.

672 *Metabolomics*

673 Untargeted fecal metabolomics was performed at the Metabolomics Research Facility of the
674 University of Calgary. Stool samples from timepoints 1, 3, 4 and 5 (N=209) were prepared for
675 metabolomic analysis. Frozen fecal samples were mixed with ice-cold 50% methanol in a 1:5
676 ratio and homogenized in a bead beater with three small steel beads (30Hz for 2x1.5 minute)
677 using high quality 2mL autoclaved safe-lock tubes. Samples were incubated for 30 min at 4 ºC
678 and then centrifuged for 10 min at maximum speed at 4 ºC. The supernatant was collected and
679 stored at -80 ºC until analysis. 200uL of each sample were transferred to 0.8mL deep 96-well
680 plates. Prior to the run samples were diluted further to 1:50. Samples were run on a Q Exactive™
681 HF Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (Thermo-Fisher, Catalog number:
682 IQLAAEGAAPFALGMBFZ) coupled to a Vanquish™ UHPLC System Integrated
683 biocompatible system (Thermo-Fisher, Catalog number: IQLAAAGABHFAPUMZZZ [8]).
684 Chromatographic separation was achieved on a Syncronis HILIC UHPLC column (2.1mm ×
685 100mm × 1.7µm, Thermo-Fisher) using a binary solvent system at a flow rate of 600uL/min.
686 Solvent A consisted of 20mM ammonium formate pH 3.0 in mass spectrometry grade H_2O;
687 Solvent B, mass spectrometry grade acetonitrile with 0.1% formic acid (%v/v). The following
688 gradients were used: 0-2 mins, 100% B; 2-7 mins, 100-80% B; 7-10 mins, 80-5% B; 10-12 mins,
689 5% B; 12-13 mins, 5-100% B; 13-15 mins, 100% B. A sample injection volume of 2µL was

used. The mass spectrometer was run in negative full scan mode at a resolution of 240,000 scanning from 50-750m/z. Metabolite data were analyzed using the MAVEN software packages[64,65]. Metabolites were identified by matching observed m/z signals (+/- 10ppm) and chromatographic retention times to those observed from commercial metabolite standards (Sigma). Creatine was quantified by an 8-point standard curve. Metabolomic data were normalized by median, square root transformed, and pareto scaled (mean-centered and divided by the square root of the standard deviation of each variable) using Metaboanalyst 5.0[66] for downstream analysis.

Immune factor determination

Frozen stool samples were used to measure cytokine, chemokine and calprotectin concentrations using the V-PLEX TH17 Panel 1, V-PLEX Proinflammatory Cytokine Panel 1, and R-PLEX Human Calprotectin assays (Mesoscale Devices). Prior to assay determination, 50 – 150 mg of sample were homogenized in 1 mL of lysis buffer (150 mM NaCl, 20 mM Tris, 1 mM EGTA, 1% Triton X-100, protease inhibitor) for 4 min at 20 Hz using a tissue homogenizer (TissueLyser II, Qiagen). Homogenized samples were then centrifuged at 14,000 x g for 10 min to removed debris, and appropriately diluted according to total protein present in corresponding supernatants, as determined by the Pierce BCA Protein Assay Kit (Thermo Scientific, Product No. 23225). Acquired MSD data for each sample was then normalized to its total protein concentration prior to statistical analysis.

Quantification and statistical analysis

Sequencing processing

Sequences were checked for quality, trimmed, merged, and checked for chimeras using the *DADA2* v1.10.1[67] pipelines for 16S or ITS2. Unique amplicon sequence variants (ASVs) were assigned taxonomy using the UNITE v.8.0 (fungi)[68] and SILVA v.132 (bacteria) [69] databases at 99% sequence similarity. Sequencing data analysis was conducted in R[70]. Initial preprocessing of the ASV table was conducted using the *Phyloseq* package v.1.26.1[71]. Overall, 10,915 unique bacterial ASVs were detected. ASVs only present in the negative controls (n=3,963) and ASVs belonging to phylum Cyanobacteria, family of mitochondria, and class of chloroplast (n=49) were removed. Samples with less than 5,000 sequencing reads were excluded (n=15) and ASVs with less than 20 reads across the entire dataset (n=6,173) were also removed. The remaining samples (n=264) were rarefied to the minimum 6,000 sequencing reads per sample resulting in 3,410 remaining ASVs. This dataset was used for analysis unless otherwise specified. For the ITS2 dataset, 3,400 unique ASVs were detected. ASVs only present in the negative controls (n=29) and ASVs belonging to kingdom Plantae (n=53) and unclassified fungi at phylum level (n=815) were removed. Samples with less than 5,000 sequencing reads were excluded (n=15) and ASVs with less than 20 reads across the entire dataset were also removed resulting in 2,319 remaining ASVs. This dataset was used for analysis unless otherwise specified.

Assessing sequencing technical accuracy

Genomic DNA of 6-8 samples was included in sequencing library preparation of both sequencing runs as biological controls. We assessed the technical accuracy between the runs by analyzing biological controls composition between the runs (**Extended Data Figure 1A and 1B**). Depth of sequencing was also compared between the sequencing runs. Run 2 had significantly higher sequencing depth per sample in both 16S rRNA and ITS2 gene sequencing

(**Extended Data Figure 1C and 1D**). Run 2 included a higher proportion of older infants and had higher total DNA concentration (**Extended Data Figure 1E and 1F**). No other variables differed between sequencing runs.

Exclusion of data

Two infants received probiotics prior to sample collection and thus their T1 samples were removed from the analysis.

Probiotic strain colonization assessment

Data analysis was conducted in R v.4.0.3[70]. The effect of the probiotic intervention and sampling timepoint on probiotic strain cell number was determined using linear mixed models (LMM) and post estimation for linear combination of coefficients using *lme4* v.1.1.26[72], *foreign* v.0.8.80 and *multcomp* v.1.4.16[73] packages. The frequency of probiotic strains detection at different timepoints were compared between controls and infants who received the probiotic using $\chi 2$ test.

Identification of microbiome community types

Microbiome maturation was assessed using hierarchical clustering on Bray-Curtis dissimilarity matrix at the genus level, with ward sum-of-square algorithm. The optimal number of clusters was determined using Gap statistics, which compares the observed change in within-cluster dispersion versus the expected change under an appropriate reference null distribution[74]. Dissimilarity (β diversity) of clusters was assessed by permutational ANOVA (PERMANOVA) using the *vegan* package v.2.5.7[75].

Assessment of the effect of probiotics on the transition to the mature community type

Markov chain state transition probabilities were estimated using *markovchain* package v.0.8.5[76] and visualized using *DiagrammeR* v.1.0.6.1[77]. The time to transition to the mature community type was assessed using Kaplan Meyer analysis using *survival* package v.3.2.7 and visualized by

survminer package v.0.4.8[78,79]. The confounding effect of other relevant early life events on the association of probiotics with gut microbiome maturation was assessed using logistic regression using *finalfit* package v.1.0.2[80].

Comparison of microbiome composition in preterm with term infants

Comparison of preterm and term infant gut microbiome was performed using the gut microbiome data of a preliminary subset of term infants enrolled in the MAGIC Study (ClinicalTrials.gov Identifier: NCT03001167), a longitudinal microbiome study of term infants conducted at the Children's Hospital of Philadelphia. We focused on breastfed, vaginally-born term infants at 1 week (N=44) and 6 months (N=24) of age. Clustering as explained above was applied to the term infant data at the genus level and compositional dissimilarity was assessed using Bray-Curtis metric and visualized using *ggridges* package v. 0.5.2[81]. The difference in PCoA1 was calculated for the preterm infants to the mean of PCoA1 of terms infants at 1 week and compared based on the intervention using ANOVA. The microbiome composition at the genus level was z normalized and visualized in a heatmap using *ComplexHeatmap* package v. 2.4.2[82].

Ecological investigation of microbiome community in response to probiotics

Microbiome network analysis was conducted at the genus level and separately for each cluster. Genera with less that 0.1% mean relative abundance and less than 25% prevalence were excluded. The microbiome data was centre log-ratio transformed to control for compositionality[83,84]. Subsequently, partial correlations were assessed using Spearman rank correlation and correlations with absolute coefficient of more than 0.25 were visualized as networks using *qgraph* package v. 1.6.5[85]. Centrality network parameters were estimated using *qgraph* package[85].

Metabolomics comparison by intervention and community type

Differential metabolic features were identified using MetaboAnalyst 5.0 with volcano plot, using a fold change threshold of 2 and adjusted t-test threshold of 0.05[66].

Predictive modelling

Predictive modelling was conducted to identify predictors of microbiome maturation in premature infants. Decision tree was performed using *rpart* v. 4.1.15 and visualized using *rpart.plot* v. 3.0.8[86,87]. Random forest was performed using 10-fold cross-validation, 500 trees, and 1000 permutation using *randomForest* v. 4.6.14 and *caret* v. 6.0.86 packages[88,89].

Structural equation modelling

Structural equation modeling (SEM) was performed using the *lavaan* package v. 0.6.6[90]. The model was estimated using maximum likelihood (ML) parameter estimation and NLMINB optimization method with bootstrapping (n=1000)[91]. Model fit was assessed by $\chi 2$ test, the comparative fix index (CFI), root mean square error of approximation (RSMEA) and its 90% confidence interval (CI), and the standardized root mean residuals (SRMR). Non-significant $\chi 2$ test, CFI$\geq$0.9, RMSEA<0.05, and SRMR<0.08 were considered as indications of good model fit [91].

Longitudinal analysis

Longitudinal analysis was performed using permuspliner function from splinectomeR v.0.1.0 with 1000 permutations[92] for taxa, and generalized estimating equation (GEE)[93] for cytokines using geepack v.1.3.2[94]. The optimum GEE model for each cytokine was selected based on the cytokine distribution and the model performance with different correlation structures: independence, exchangeable, autoregressive 1, or unstructured. The family of the GEE model was set as gaussian or gamma for normal or positively skewed cytokine distribution,

respectively. The models were compared based on the quasi-likelihood information (QIC) criterion using MuMIn v.1.43.17 and pander v.0.6.4 packages [95,96]. The model with the lowest QIC was selected for each cytokine. Trend analysis was conducted using trendyspliner function of SplinectomeR.

Univariate analysis of cytokines and metabolites

Cytokine and metabolite concentrations were compared by pairwise Wilcoxon test.

Data Availability Statement

Demultiplexed 16S and ITS2 sequencing data was deposited into the Sequence Read Archive (SRA) of NCBI and will be accessible via accession numbers PRJNA721684 and PRJNA721688. Metabolomics mass spectral raw data were deposited to MetaboLights (study identifier MTBLS2699).

Code Availability Statement: The R codes are provided as supplementary file 1.

References

1 Koenig, J. E. *et al.* Succession of microbial consortia in the developing infant gut microbiome. *Proc Natl Acad Sci U S A* **108 Suppl 1**, 4578-4585, doi:10.1073/pnas.1000081107 (2011).

2 Yassour, M. *et al.* Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability. *Sci Transl Med* **8**, 343ra381, doi:10.1126/scitranslmed.aad0917 (2016).

3 Rao, C. *et al.* Multi-kingdom ecological drivers of microbiota assembly in preterm infants. *Nature* **591**, 633-638, doi:10.1038/s41586-021-03241-8 (2021).

4 Mitchell, C. M. *et al.* Delivery mode affects stability of early infant gut microbiota. *Cell Rep Med* **1**, 100156, doi:10.1016/j.xcrm.2020.100156 (2020).

5 Stewart, C. J. *et al.* Temporal bacterial and metabolic development of the preterm gut reveals specific signatures in health and disease. *Microbiome* **4**, 67, doi:10.1186/s40168-016-0216-8 (2016).

6 La Rosa, P. S. *et al.* Patterned progression of bacterial populations in the premature infant gut. *Proc Natl Acad Sci U S A* **111**, 12522-12527, doi:10.1073/pnas.1409497111 (2014).

7 Korpela, K. *et al.* Intestinal microbiota development and gestational age in preterm neonates. *Sci Rep* **8**, 2453, doi:10.1038/s41598-018-20827-x (2018).

8 Chernikova, D. A. *et al.* The premature infant gut microbiome during the first 6 weeks of life differs based on gestational maturity at birth. *Pediatr Res* **84**, 71-79, doi:10.1038/s41390-018-0022-z (2018).

9 Wandro, S. *et al.* The microbiome and metabolome of preterm infant stool are personalized and not driven by health outcomes, including necrotizing enterocolitis and late-onset sepsis. *mSphere* **3**, e00104-00118, doi:10.1128/mSphere.00104-18 (2018).

10 Olin, A. *et al.* Stereotypic immune system development in newborn children. *Cell* **174**, 1277-1292 e1214, doi:10.1016/j.cell.2018.06.045 (2018).

11 Arboleya, S. *et al.* Establishment and development of intestinal microbiota in preterm neonates. *FEMS Microbiol Ecol* **79**, 763-772, doi:10.1111/j.1574-6941.2011.01261.x (2012).

12 Costello, E. K., Carlisle, E. M., Bik, E. M., Morowitz, M. J. & Relman, D. A. Microbiome assembly across multiple body sites in low-birthweight infants. *mBio* **4**, e00782-00713, doi:10.1128/mBio.00782-13 (2013).

13 Henderickx, J. G. E., Zwittink, R. D., van Lingen, R. A., Knol, J. & Belzer, C. The preterm gut microbiota: an inconspicuous challenge in nutritional neonatal care. *Front Cell Infect Microbiol* **9**, 85, doi:10.3389/fcimb.2019.00085 (2019).

14 Laforest-Lapointe, I. & Arrieta, M. C. Patterns of early-life gut microbial colonization during human immune development: an ecological perspective. *Front Immunol* **8**, 788, doi:10.3389/fimmu.2017.00788 (2017).

15 Pammi, M. *et al.* Intestinal dysbiosis in preterm infants preceding necrotizing enterocolitis: a systematic review and meta-analysis. *Microbiome* **5**, 31, doi:10.1186/s40168-017-0248-8 (2017).

16 Masi, A. C. & Stewart, C. J. The role of the preterm intestinal microbiome in sepsis and necrotising enterocolitis. *Early Hum Dev* **138**, 104854, doi:10.1016/j.earlhumdev.2019.104854 (2019).
17 Vongbhavit, K. & Underwood, M. A. Prevention of necrotizing enterocolitis through manipulation of the intestinal microbiota of the premature infant. *Clin Ther* **38**, 716-732, doi:10.1016/j.clinthera.2016.01.006 (2016).
18 Bertelsen, R. J., Jensen, E. T. & Ringel-Kulka, T. Use of probiotics and prebiotics in infant feeding. *Best Pract Res Clin Gastroenterol* **30**, 39-48, doi:10.1016/j.bpg.2016.01.001 (2016).
19 Aceti, A. *et al.* Probiotics prevent late-onset sepsis in human milk-fed, very low birth weight preterm infants: systematic review and meta-analysis. *Nutrients* **9**, doi:10.3390/nu9080904 (2017).
20 Alcon-Giner, C. *et al.* Microbiota supplementation with Bifidobacterium and Lactobacillus modifies the preterm infant gut microbiota and metabolome: an observational study. *Cell Rep Med* **1**, 100077, doi:10.1016/j.xcrm.2020.100077 (2020).
21 Plummer, E. L. *et al.* Gut microbiota of preterm infants supplemented with probiotics: sub-study of the ProPrems trial. *BMC Microbiol* **18**, 184, doi:10.1186/s12866-018-1326-1 (2018).
22 Brussow, H. Problems with the concept of gut microbiota dysbiosis. *Microb Biotechnol* **13**, 423-434, doi:10.1111/1751-7915.13479 (2020).
23 Ofek Shlomai, N., Deshpande, G., Rao, S. & Patole, S. Probiotics for preterm neonates: what will it take to change clinical practice? *Neonatology* **105**, 64-70, doi:10.1159/000354891 (2014).
24 Modi, N. Probiotics and necrotising enterocolitis: the devil (as always) is in the detail. *Neonatology* **105**, 71-73, doi:10.1159/000354909 (2014).
25 Frese, S. A. *et al.* Persistence of Supplemented Bifidobacterium longum subsp. infantis EVC001 in Breastfed Infants. *mSphere* **2**, doi:10.1128/mSphere.00501-17 (2017).
26 Henrick, B. M. *et al.* Bifidobacteria-mediated immune system imprinting early in life. *Cell* **184**, 3884-3898 e3811, doi:10.1016/j.cell.2021.05.030 (2021).
27 Egan, M. *et al.* Cross-feeding by Bifidobacterium breve UCC2003 during co-cultivation with Bifidobacterium bifidum PRL2010 in a mucin-based medium. *BMC Microbiol* **14**, 282, doi:10.1186/s12866-014-0282-7 (2014).
28 Cheng, C. C. *et al.* Ecological Importance of Cross-Feeding of the Intermediate Metabolite 1,2-Propanediol between Bacterial Gut Symbionts. *Appl Environ Microbiol* **86**, doi:10.1128/AEM.00190-20 (2020).
29 Tannock, G. W. Building Robust Assemblages of Bacteria in the Human Gut in Early Life. *Appl Environ Microbiol* **87**, e0144921, doi:10.1128/AEM.01449-21 (2021).
30 Connell, J. H. & Slatyer, R. H. Mechanisms of succession in natural communities and their role in community stability and organization. *Amer. Natur.* **111**, 1119-1144 (1977).
31 Young, T. P., Petersen, D. A. & Clary, J. J. The ecology of restoration: historical links, emerging issues and unexplored realms. *Ecology Letters* **8**, 662-673, doi:10.1111/j.1461-0248.2005.00764.x (2005).

32 Ridlon, J. M., Harris, S. C., Bhowmik, S., Kang, D. J. & Hylemon, P. B. Consequences of bile salt biotransformations by intestinal bacteria. *Gut Microbes* **7**, 22-39, doi:10.1080/19490976.2015.1127483 (2016).
33 Boehm, G., Braun, W., Moro, G. & Minoli, I. Bile acid concentrations in serum and duodenal aspirates of healthy preterm infants: effects of gestational and postnatal age. *Biol Neonate* **71**, 207-214, doi:10.1159/000244419 (1997).
34 Ramiro-Cortijo, D. *et al.* Breast Milk Lipids and Fatty Acids in Regulating Neonatal Intestinal Development and Protecting against Intestinal Injury. *Nutrients* **12**, doi:10.3390/nu12020534 (2020).
35 Murata, M. & Kawanishi, S. Oxidative DNA damage induced by nitrotyrosine, a biomarker of inflammation. *Biochem Biophys Res Commun* **316**, 123-128, doi:10.1016/j.bbrc.2004.02.022 (2004).
36 Banks, B. A., Ischiropoulos, H., McClelland, M., Ballard, P. L. & Ballard, R. A. Plasma 3-nitrotyrosine is elevated in premature infants who develop bronchopulmonary dysplasia. *Pediatrics* **101**, 870-874, doi:10.1542/peds.101.5.870 (1998).
37 Sheffield, M., Mabry, S., Thibeault, D. W. & Truog, W. E. Pulmonary nitric oxide synthases and nitrotyrosine: findings during lung development and in chronic lung disease of prematurity. *Pediatrics* **118**, 1056-1064, doi:10.1542/peds.2006-0195 (2006).
38 Egan, C. E. *et al.* Toll-like receptor 4-mediated lymphocyte influx induces neonatal necrotizing enterocolitis. *J Clin Invest* **126**, 495-508, doi:10.1172/JCI83356 (2016).
39 Ferrario, C. *et al.* Exploring Amino Acid Auxotrophy in Bifidobacterium bifidum PRL2010. *Front Microbiol* **6**, 1331, doi:10.3389/fmicb.2015.01331 (2015).
40 Chow, J. *et al.* Fecal metabolomics of healthy breast-fed versus formula-fed infants before and during in vitro batch culture fermentation. *J Proteome Res* **13**, 2534-2542, doi:10.1021/pr500011w (2014).
41 Li, N. *et al.* Distinct Gut Microbiota and Metabolite Profiles Induced by Different Feeding Methods in Healthy Chinese Infants. *Front Microbiol* **11**, 714, doi:10.3389/fmicb.2020.00714 (2020).
42 Martinez, I. *et al.* Experimental evaluation of the importance of colonization history in early-life gut microbiota assembly. *Elife* **7**, e36521, doi:10.7554/eLife.36521 (2018).
43 Hu, H. J., Zhang, G. Q., Zhang, Q., Shakya, S. & Li, Z. Y. Probiotics prevent Candida colonization and invasive fungal sepsis in preterm neonates: a systematic review and meta-Analysis of randomized controlled trials. *Pediatr Neonatol* **58**, 103-110, doi:10.1016/j.pedneo.2016.06.001 (2017).
44 Manzoni, P. *et al.* Oral supplementation with Lactobacillus casei subspecies rhamnosus prevents enteric colonization by Candida species in preterm neonates: a randomized study. *Clin Infect Dis* **42**, 1735-1742, doi:10.1086/504324 (2006).
45 Saiki, T., Mitsuyama, K., Toyonaga, A., Ishida, H. & Tanikawa, K. Detection of pro- and anti-inflammatory cytokines in stools of patients with inflammatory bowel disease. *Scand J Gastroenterol* **33**, 616-622, doi:10.1080/00365529850171891 (1998).
46 Dudakov, J. A., Hanash, A. M. & van den Brink, M. R. Interleukin-22: immunobiology and pathology. *Annu Rev Immunol* **33**, 747-785, doi:10.1146/annurev-immunol-032414-112123 (2015).

47 Mihi, B. *et al.* Interleukin-22 signaling attenuates necrotizing enterocolitis by promoting epithelial cell regeneration. *Cell Rep Med* **2**, 100320, doi:10.1016/j.xcrm.2021.100320 (2021).

48 Connell, J. H. S. R. O. Mechanisms of succession in natural communities and their role in community stability and organization. *The American Naturalist* **111**, doi:10.1086/283241 (1977).

49 Lee, J. Y., Sohn, K. H., Rhee, S. H. & Hwang, D. Saturated fatty acids, but not unsaturated fatty acids, induce the expression of cyclooxygenase-2 mediated through Toll-like receptor 4. *J Biol Chem* **276**, 16683-16689, doi:10.1074/jbc.M011695200 (2001).

50 Cho, S. X., Berger, P. J., Nold-Petry, C. A. & Nold, M. F. The immunological landscape in necrotising enterocolitis. *Expert Rev Mol Med* **18**, e12, doi:10.1017/erm.2016.13 (2016).

51 Maheshwari, A. *et al.* Cytokines associated with necrotizing enterocolitis in extremely-low-birth-weight infants. *Pediatr Res* **76**, 100-108, doi:10.1038/pr.2014.48 (2014).

52 O'Callaghan, A. & van Sinderen, D. Bifidobacteria and their role as members of the human gut microbiota. *Front Microbiol* **7**, 925, doi:10.3389/fmicb.2016.00925 (2016).

53 Walter, J., Maldonado-Gomez, M. X. & Martinez, I. To engraft or not to engraft: an ecological framework for gut microbiome modulation with live microbes. *Curr Opin Biotechnol* **49**, 129-139, doi:10.1016/j.copbio.2017.08.008 (2018).

54 Fukuda, S. *et al.* Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* **469**, 543-547, doi:10.1038/nature09646 (2011).

55 Falony, G., Vlachou, A., Verbrugghe, K. & De Vuyst, L. Cross-feeding between Bifidobacterium longum BB536 and acetate-converting, butyrate-producing colon bacteria during growth on oligofructose. *Appl Environ Microbiol* **72**, 7835-7841, doi:10.1128/AEM.01296-06 (2006).

56 Belenguer, A. *et al.* Two routes of metabolic cross-feeding between Bifidobacterium adolescentis and butyrate-producing anaerobes from the human gut. *Appl Environ Microbiol* **72**, 3593-3599, doi:10.1128/AEM.72.5.3593-3599.2006 (2006).

57 Centanni, M., Ferguson, S. A., Sims, I. M., Biswas, A. & Tannock, G. W. Bifidobacterium bifidum ATCC 15696 and Bifidobacterium breve 24b Metabolic Interaction Based on 2'-O-Fucosyl-Lactose Studied in Steady-State Cultures in a Freter-Style Chemostat. *Appl Environ Microbiol* **85**, doi:10.1128/AEM.02783-18 (2019).

58 Gotoh, A. *et al.* Sharing of human milk oligosaccharides degradants within bifidobacterial communities in faecal cultures supplemented with Bifidobacterium bifidum. *Sci Rep* **8**, 13958, doi:10.1038/s41598-018-32080-3 (2018).

59 Sprockett, D., Fukami, T. & Relman, D. A. Role of priority effects in the early-life assembly of the gut microbiota. *Nat Rev Gastroenterol Hepatol* **15**, 197-205, doi:10.1038/nrgastro.2017.173 (2018).

60 Walter, J. & Ley, R. The human gut microbiome: ecology and recent evolutionary changes. *Annu Rev Microbiol* **65**, 411-429, doi:10.1146/annurev-micro-090110-102830 (2011).

61 Zheng, J. *et al.* A taxonomic note on the genus Lactobacillus: Description of 23 novel genera, emended description of the genus Lactobacillus Beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. *Int J Syst Evol Microbiol* **70**, 2782-2858, doi:10.1099/ijsem.0.004107 (2020).

62 Ford, A. L. *et al.* Microbiota stability and gastrointestinal tolerance in response to a high-protein diet with and without a prebiotic, probiotic, and synbiotic: a randomized, double-blind, placebo-controlled trial in older women. *J Acad Nutr Diet* **120**, 500-516 e510, doi:10.1016/j.jand.2019.12.009 (2020).
63 Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* **79**, 5112-5120, doi:10.1128/AEM.01043-13 (2013).
64 Clasquin, M. F., Melamud, E. & Rabinowitz, J. D. LC-MS data processing with MAVEN: a metabolomic analysis and visualization engine. *Curr Protoc Bioinformatics* **Chapter 14**, Unit14 11, doi:10.1002/0471250953.bi1411s37 (2012).
65 Melamud, E., Vastag, L. & Rabinowitz, J. D. Metabolomic analysis and visualization engine for LC-MS data. *Anal Chem* **82**, 9818-9826, doi:10.1021/ac1021166 (2010).
66 Pang, Z. *et al.* MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res* **49**, W388-W396, doi:10.1093/nar/gkab382 (2021).
67 Callahan, B. J. *et al.* DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581-583, doi:10.1038/nmeth.3869 (2016).
68 Nilsson, R. H. *et al.* The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res* **47**, D259-D264, doi:10.1093/nar/gky1022 (2019).
69 Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**, D590-596, doi:10.1093/nar/gks1219 (2013).
70 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2017).
71 McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217, doi:10.1371/journal.pone.0061217 (2013).
72 Bates, D., Maechler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. . *J Stat Softw* **67**, 1-48 (2015).
73 Hothorn, T., Bretz, f. & Westfall, p. Simultaneous inference in general parametric models. *Biom J* **50**, 346-363 (2008).
74 Tibshirani, R., Walther, G. & Hastie, T. Estimating the number of clusters in a data set via the gap statistic. *J R Statist Soc B* **63**, 411-423 (2001).
75 Oksanen, J. *et al.* Vegan: community ecology package. R package version 2.4-3. . (2017).
76 Spedicato, G. A. Discrete time Markov Chains with R. *R J* **9**, 84-104 (2017).
77 Iannone, R. DiagrammeR: graph/network visualization. R package version 1.0.6.1. https://CRAN.R-project.org/package=DiagrammeR. (2020).
78 Therneau, T. A package for survival analysis in R. R package version 3.2-7. https://CRAN.R-project.org/package=survival. (2020).
79 Kassambara, A., Kosinski, M. & Biecek, P. survminer: drawing survival curves using 'ggplot2'. R package version 0.4.8. https://CRAN.R-project.org/package=survminer. (2020).

80 Harrison, E., Drake, T. & Ots, R. finalfit: quickly create elegant regression results tables and plots when modelling. R package version 1.0.2. https://CRAN.R-project.org/package=finalfit. (2020).
81 Wilke, C. O. ggridges: ridgeline plots in 'ggplot2'. R package version 0.5.2. https://CRAN.R-project.org/package=ggridges. (2020).
82 Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847-2849, doi:10.1093/bioinformatics/btw313 (2016).
83 Palarea-Albaladejo, J. & Martin-Fernandez, J. A. zCompositions -- R package for multivariate imputation of left-censored data under a compositional approach. *Chemom Intell Lab Syst* **143**, 85-96 (2015).
84 Gloor, G. B. & Reid, G. Compositional analysis: a valid approach to analyze microbiome high-throughput sequencing data. *Can J Microbiol* **62**, 692-703, doi:10.1139/cjm-2015-0821 (2016).
85 Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D. & Borsboom, D. qgraph: network visualizations of relationships in psychometric data. *J Stat Softw* **48**, 1-18 (2012).
86 Therneau, T. & Atkinson, B. rpart: recursive partitioning and regression trees. R package version 4.1-15. https://CRAN.R-project.org/package=rpart. (2019).
87 Milborrow, S. rpart.plot: plot 'rpart' models: an enhanced version of 'plot.rpart'. R package version 3.0.8. https://CRAN.R-project.org/package=rpart.plot. (2019).
88 Liaw, A. & Wiener, M. Classification and regression by randomForest *R News* **2**, 18-22 (2002).
89 Kuhn, M. caret: classification and regression training. R package version 6.0-86. https://CRAN.R-project.org/package=caret. (2020).
90 Rosseel, Y. lavaan: an R package for structural equation modeling. *J Stat Softw* **48**, 1-36 (2012).
91 Kline, R. B. *Principles and practices of structural equation modeling*. 4 edn, (The Guilford Press, 2016).
92 Shields-Cutler, R. R., Al-Ghalith, G. A., Yassour, M. & Knights, D. SplinectomeR enables group comparisons in longitudinal microbiome studies. *Front Microbiol* **9**, 785, doi:10.3389/fmicb.2018.00785 (2018).
93 Pekár, S., Brabec, M. & Bshary, R. Generalized estimating equations: A pragmatic and flexible approach to the marginal GLM modelling of correlated data in the behavioural sciences. *Ethology* **124**, 86-93, doi:10.1111/eth.12713 (2018).
94 Højsgaard, S., Halekoh, U. & Yan, J. The R Package geepack for Generalized Estimating Equations
J Stat Softw **15**, 1-11 (2006).
95 Bartoń, K. MuMIn: Multi-Model Inference. R package version 1.43.17. https://CRAN.R-project.org/package=MuMIn. (2020).
96 Daróczi, G. & Tsegelskyi, R. pander: An R 'Pandoc' Writer. R package version 0.6.3. https://CRAN.R-project.org/package=pander. (2018).

1077

1078

1079

1080

Figure 1

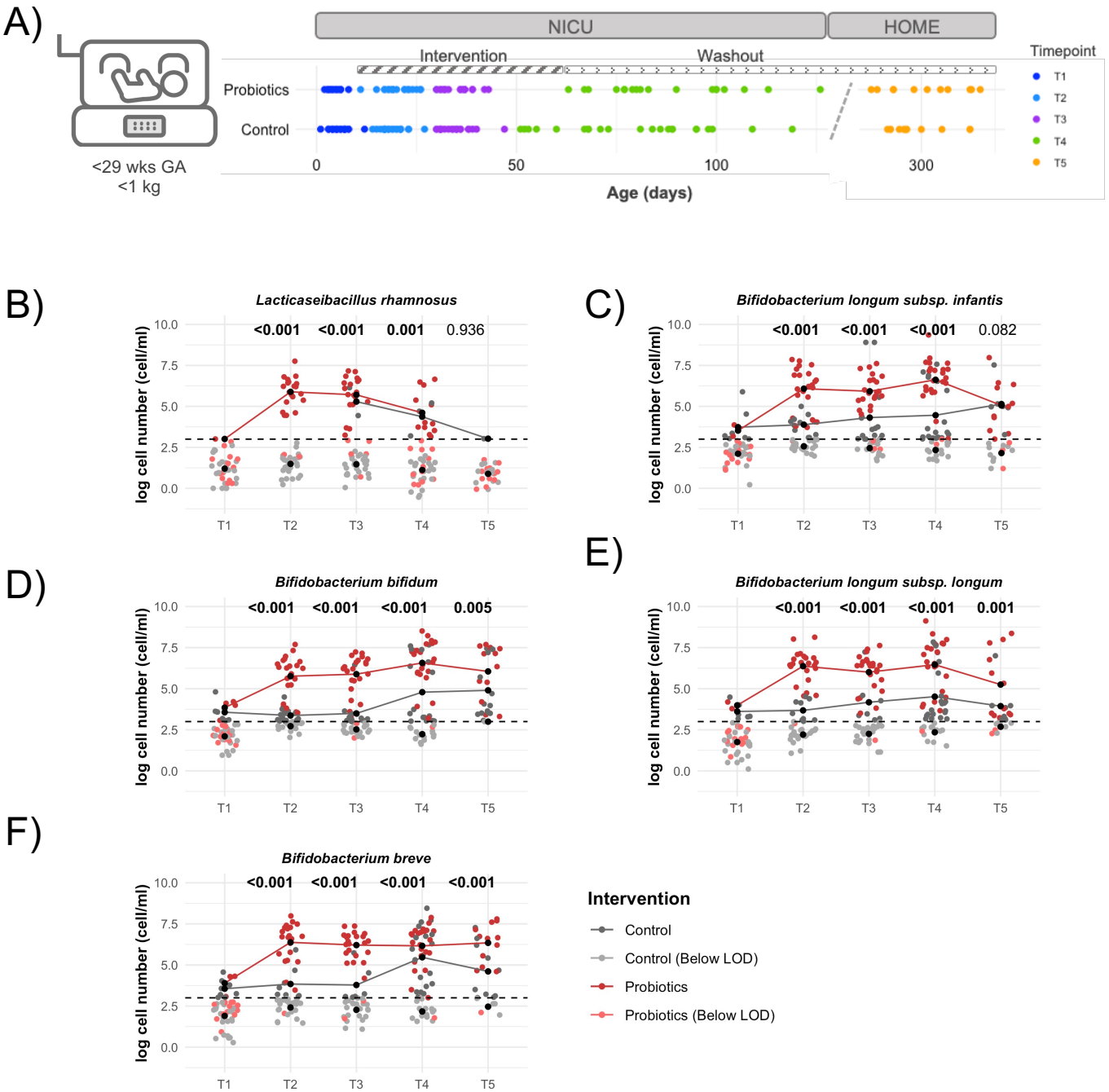


Figure 1. Probiotic strains can stably colonize the extremely premature infant gut. A) Study design for the randomized controlled trial of probiotics in extremely preterm infants. In the treatment group, probiotic was started in the first week of life before sample collection (T1) and continued until 37-39 weeks gestational age (GA) weeks spanning T2 and T3. Additional samples were collected after cessation of probiotic at 39-40 weeks GA (T4) and 6 months corrected age (CA) (T5). **B-F)** Concentration of probiotic strains assessed by strain-specific qPCR demonstrates increased concentration of all probiotics strains immediately after starting probiotic at T2. *Lacticaseibacillus rhamnosus* decreased after cessation of probiotics (B) while the *Bifidobacterium* strains showed stable colonisation until 6 months CA. The dashed line denotes the limit of detection (10^3 bacterial cells/ml). P values are obtained from linear mixed models (LMM) and post estimation for linear combination of coefficients (see also Extended Data Table 2). LOD, limit of detection.

Figure 2

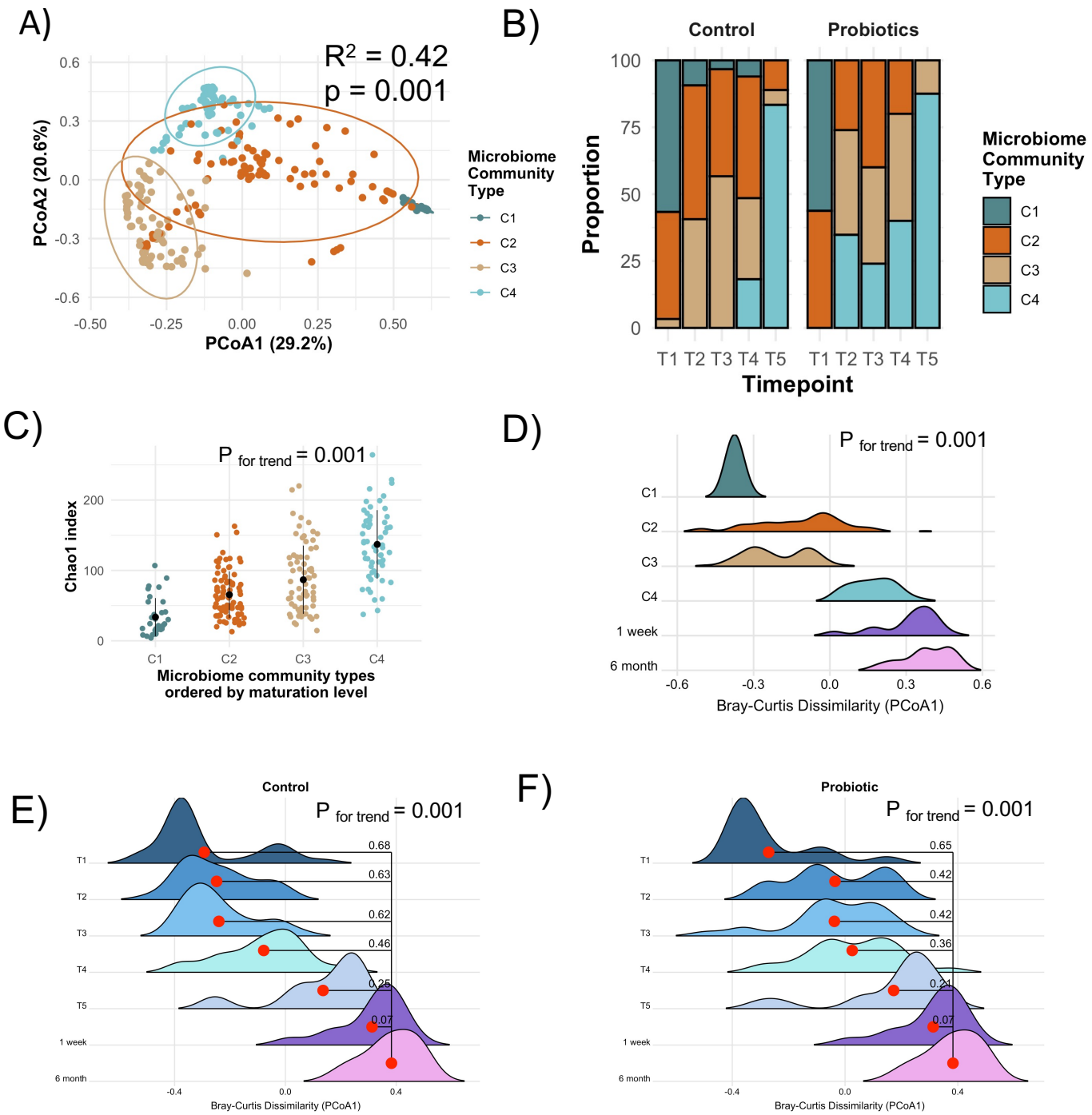


Figure 2. Probiotics accelerate gut microbiome maturation in extremely preterm infants. A) Four gut microbiome community types were identified using hierarchical clustering on Bray-Curtis dissimilarity matrix. Association of the community types with beta diversity was tested using PERMANOVA. **B)** Microbiome community type distribution across timepoints and probiotic use. Community types showed temporal distribution, with C1 and C2 more frequent in earlier and C4 in later timepoints. As a result, C4 is considered the mature community type, which appeared earlier in infants treated with probiotics. **C)** Comparison of bacterial richness (Chao1) in community types (See Extended **Figure S2D** for comparison of beta diversity). **D)** Comparison of the maturational patterns of the microbiome community types with term infants at 1 week and 6 months of age. **E-F)** Comparison of the temporal development of preterm infant microbiome with term infants at 1 week and 6 months of age in controls **(E)** and probiotic-treated infants **(F)**. Centroid of each timepoint is denoted as the red circle and the distance to the centroid of each timepoint to the centroid of 6-month term infants are presented as labels. Trend analysis in panels C-F were conducted using trendyspliner in SplinectomeR package[70].

Figure 3.

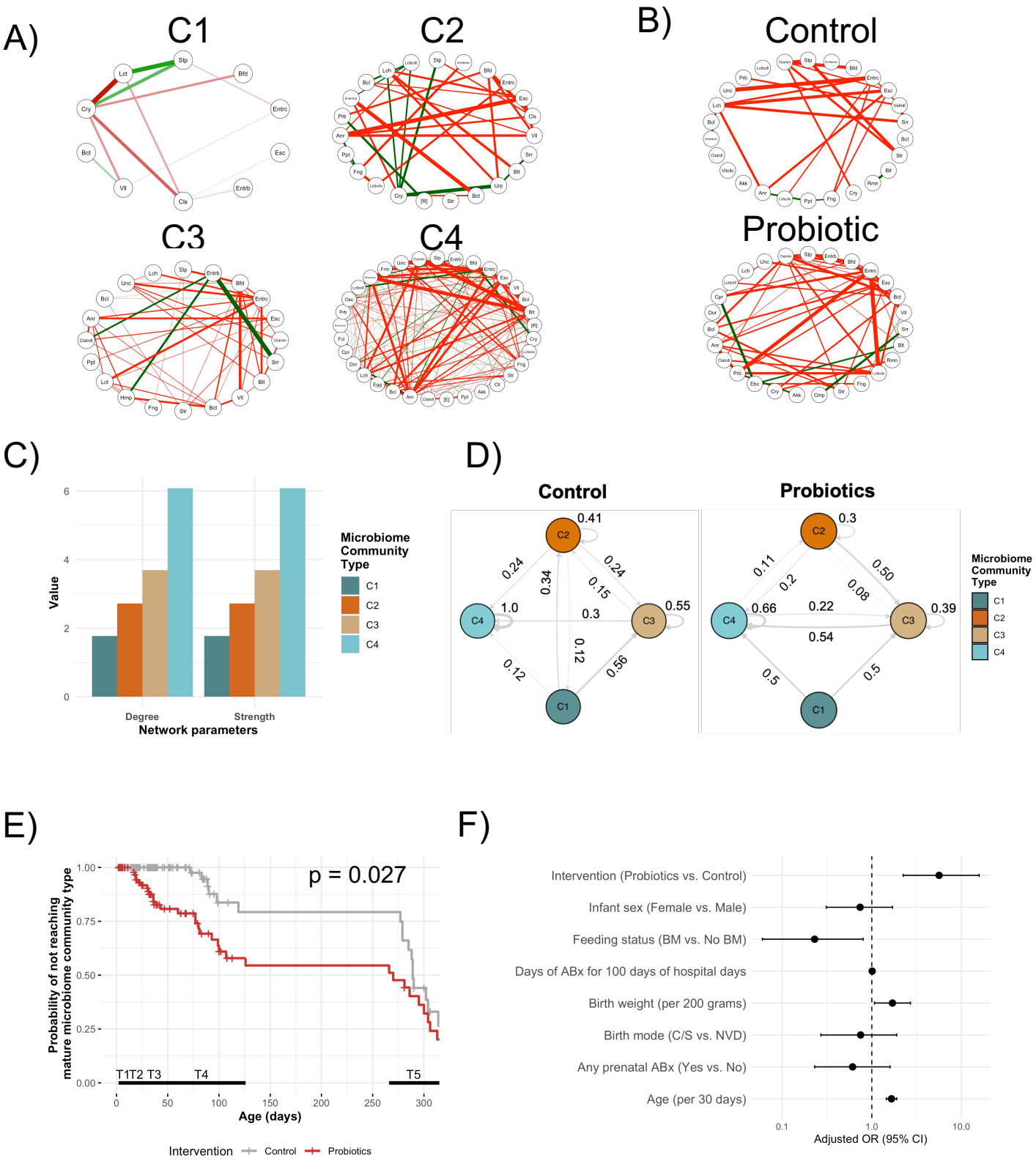


Figure 3. Probiotics promote a microbial community with higher interconnectivity and stability. A-B) Network analysis of the preterm infant microbiome along the microbiome maturation trajectory **(A)** and by intervention **(B)**. **C)** Comparison of network degree and strength across community types. **D)** Probability of transition between community types assessed by Markov Chain modelling compared in controls and probiotic group. **E)** Time-to-event analysis demonstrates that probiotics accelerates transition into the C4 mature community type. Kaplan-Meyer curve for the probability of not reaching the mature community type is shown. **F)** Multivariable logistic regression demonstrating the association of probiotic treatment with microbiome maturation independently of early life events. Adjusted Odds Ratio (OR) and 95% confidence interval (CI) are presented for all variables in the model.

Figure 4

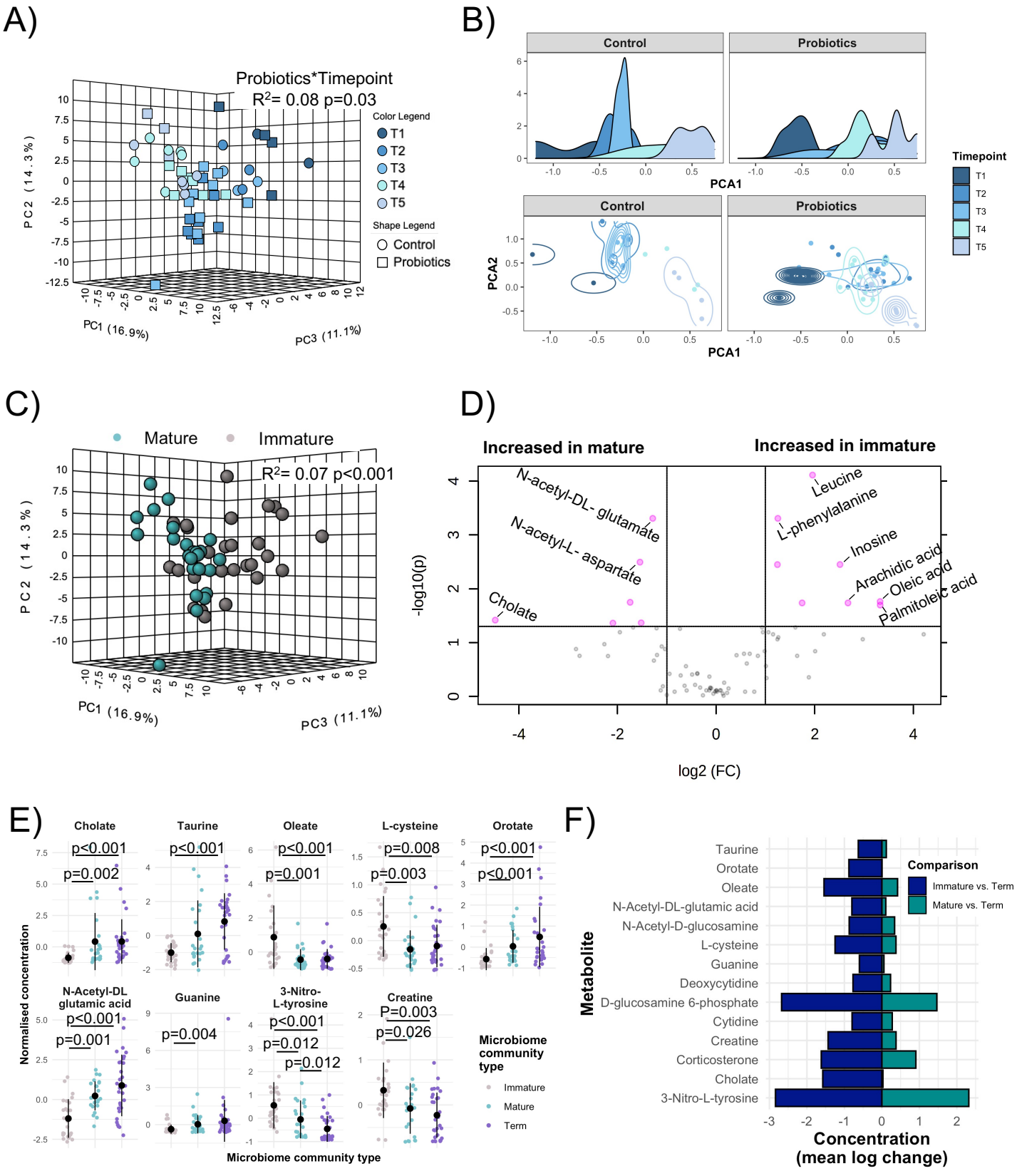

A)
Probiotics*Timepoint
R²= 0.08 p=0.03
Color Legend
T1
T2
T3
T4
T5
Shape Legend
Control
Probiotics
PC2 (14.3%)
PC1 (16.9%)
PC3 (11.1%)
B)
Control
Probiotics
PCA1
PCA2
Timepoint
T1
T2
T3
T4
T5
C)
Mature
Immature
R²= 0.07 p<0.001
PC2 (14.3%)
PC1 (16.9%)
PC3 (11.1%)
D)
Increased in mature
Increased in immature
N-acetyl-DL- glutamate
N-acetyl-L- aspartate
Cholate
Leucine
L-phenylalanine
Inosine
Arachidic acid
Oleic acid
Palmitoleic acid
-log10(p)
log2 (FC)
E)
Cholate
p<0.001
p=0.002
Taurine
p<0.001
Oleate
p<0.001
p=0.001
L-cysteine
p=0.008
p=0.003
Orotate
p<0.001
p<0.001
N-Acetyl-DL glutamic acid
p<0.001
p=0.001
Guanine
p=0.004
3-Nitro-L-tyrosine
p<0.001
p=0.012
p=0.012
Creatine
P=0.003
p=0.026
Normalised concentration
Microbiome community type
Microbiome community type
Immature
Mature
Term
F)
Taurine
Orotate
Oleate
N-Acetyl-DL-glutamic acid
N-Acetyl-D-glucosamine
L-cysteine
Guanine
Deoxycytidine
D-glucosamine 6-phosphate
Cytidine
Creatine
Corticosterone
Cholate
3-Nitro-L-tyrosine
Metabolite
Comparison
Immature vs. Term
Mature vs. Term
Concentration
(mean log change)

Figure 4. Probiotic-induced microbiome maturation is reflected in the stool metabolome. A-B) Principal component analysis of gut metabolome in premature infants at different timepoints and by intervention. Interaction between the effects of timepoint and probiotics was tested using PERMANOVA. **C)** Principal component analysis of gut metabolome in premature infants with mature (C4) vs. immature (C1-C3) community types. Effect of maturational status on the variance of the metabolome was tested using PERMANOVA. **D)** Differentially enriched metabolites in mature (C4) vs. immature (C1-C3) community types as assessed by volcano plot with fold change threshold of 2 and adjusted t-test threshold of 0.05. Pink circles represent features above this threshold. **E)** The most discriminatory metabolic features from immature (gray) or mature (turquoise) microbiome maturation status in premature infants compared to term, breastfed infants (purple). Comparisons were made by pairwise Wilcoxon test. **F)** Metabolite levels by microbiome maturity in relation to term breastfed infants. Mean fold difference in the mature-term vs. immature-term comparisons are shown .

Figure 5

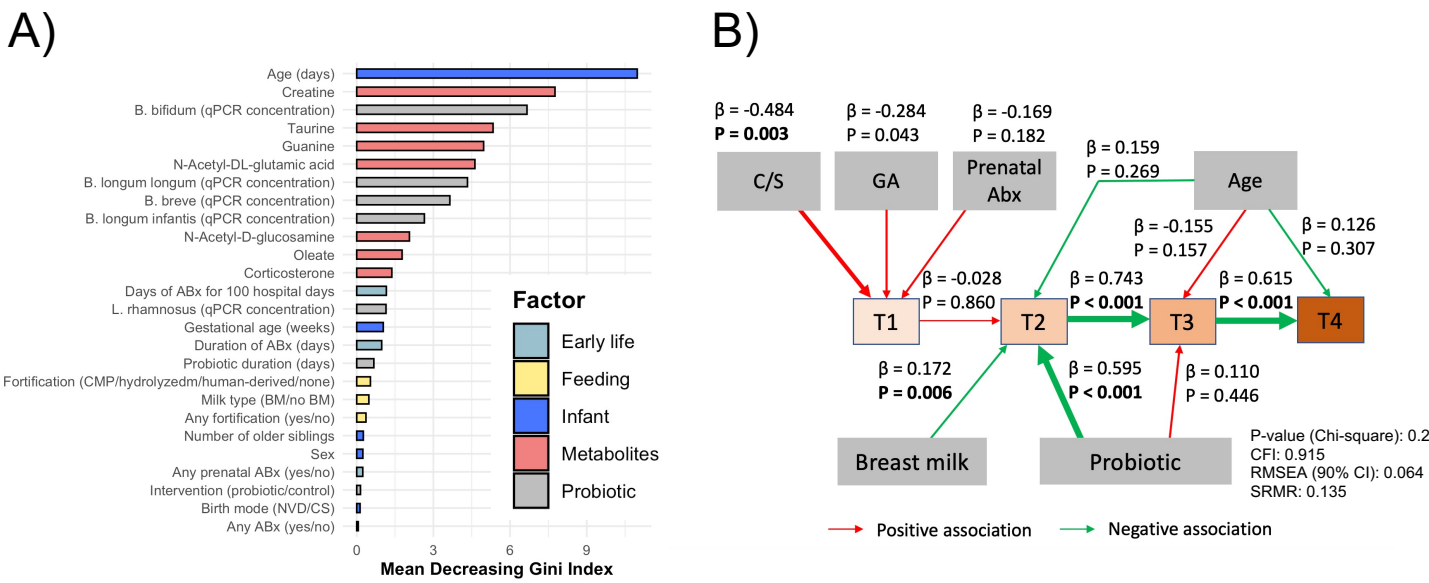


Figure 5. Probiotic strains and stool metabolites are predictive and drivers of microbiome maturation. A) Predictors of mature microbiome community type (C4 vs. C1 C2 & C3 combined) ordered by their importance identified through random forest modelling using 10-fold cross-validation, 500 trees, and 1000 permutations. **B)** Structural equation modelling was used to differentiate the influence of probiotics on bacterial richness (Chao1) at each timepoint while taking into account the structure of association of other early life factors. Probiotic was administered during T2 and T3 timepoints. Model fit was assessed using p value, CFI, RMSEA, and SRMR. Abx, antibiotics; CFI, comparative fit index; C/S, Caesarean section; RMSEA, root mean square error of approximation; SRMR, standardized root mean residuals.

Figure 6

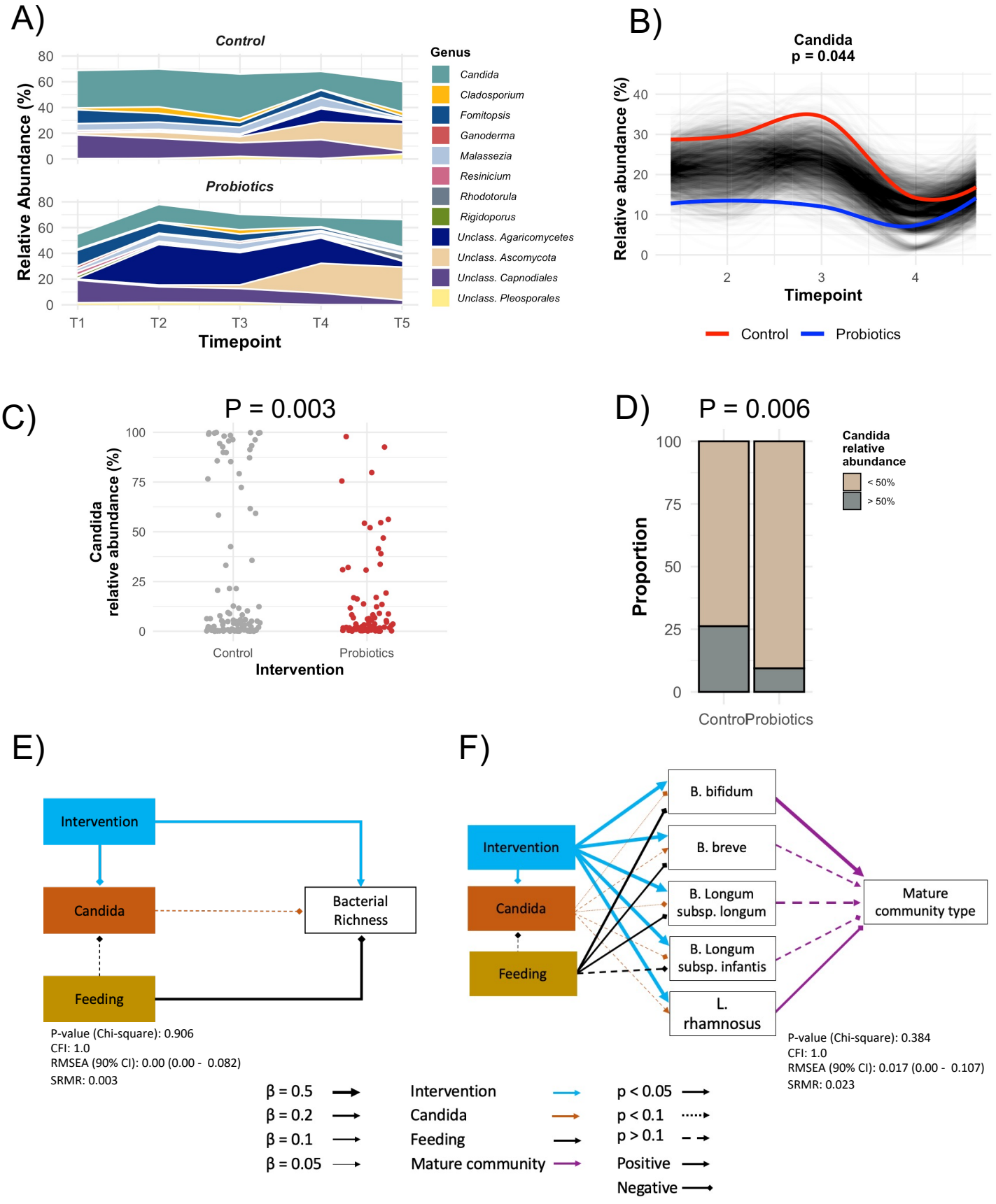

A)
Control
Probiotics
Relative Abundance (%)
Timepoint
T1
T2
T3
T4
T5
Genus
Candida
Cladosporium
Fomitopsis
Ganoderma
Malassezia
Resinicium
Rhodotorula
Rigidoporus
Unclass. Agaricomycetes
Unclass. Ascomycota
Unclass. Capnodiales
Unclass. Pleosporales
B)
Candida
p = 0.044
Relative abundance (%)
Timepoint
Control
Probiotics
C)
P = 0.003
Candida relative abundance (%)
Control
Probiotics
Intervention
D)
P = 0.006
Proportion
Candida relative abundance
< 50%
> 50%
Contro
Probiotics
E)
Intervention
Candida
Feeding
Bacterial Richness
P-value (Chi-square): 0.906
CFI: 1.0
RMSEA (90% CI): 0.00 (0.00 - 0.082)
SRMR: 0.003
F)
Intervention
Candida
Feeding
B. bifidum
B. breve
B. Longum subsp. longum
B. Longum subsp. infantis
L. rhamnosus
Mature community type
P-value (Chi-square): 0.384
CFI: 1.0
RMSEA (90% CI): 0.017 (0.00 - 0.107)
SRMR: 0.023
β = 0.5
β = 0.2
β = 0.1
β = 0.05
Intervention
Candida
Feeding
Mature community
p < 0.05
p < 0.1
p > 0.1
Positive
Negative

Figure 6. Probiotic use depletes *Candida* spp. but probiotic-*Candida* interactions do not modulate microbiome maturation. A) Mycobiome community structure at genus level compared in controls and infants who received probiotics. **B)** Longitudinal analysis of *Candida* spp. according to the intervention using splinectomeR reveals significantly lower abundance in the probiotic group. **C)** Distribution of *Candida* spp. by intervention confirms lower average relative abundance in the probiotic group. **D)** Categorizing *Candida* spp. relative abundance into < 50% or >50% revealed the infants who received probiotic are less frequently dominated by high levels of *Candida* spp. **E-F)** Structural equation modelling to examine the direct effect of *Candida* spp. on bacterial richness **E)** and indirect effect on microbiome maturation via interaction with probiotic strains **F).** Model fit was assessed using p value, CFI, RMSEA, and SRMR. CFI, comparative fit index; C/S, Caesarean section; RMSEA, root mean square error of approximation; SRMR, standardized root mean residuals.

Figure 7

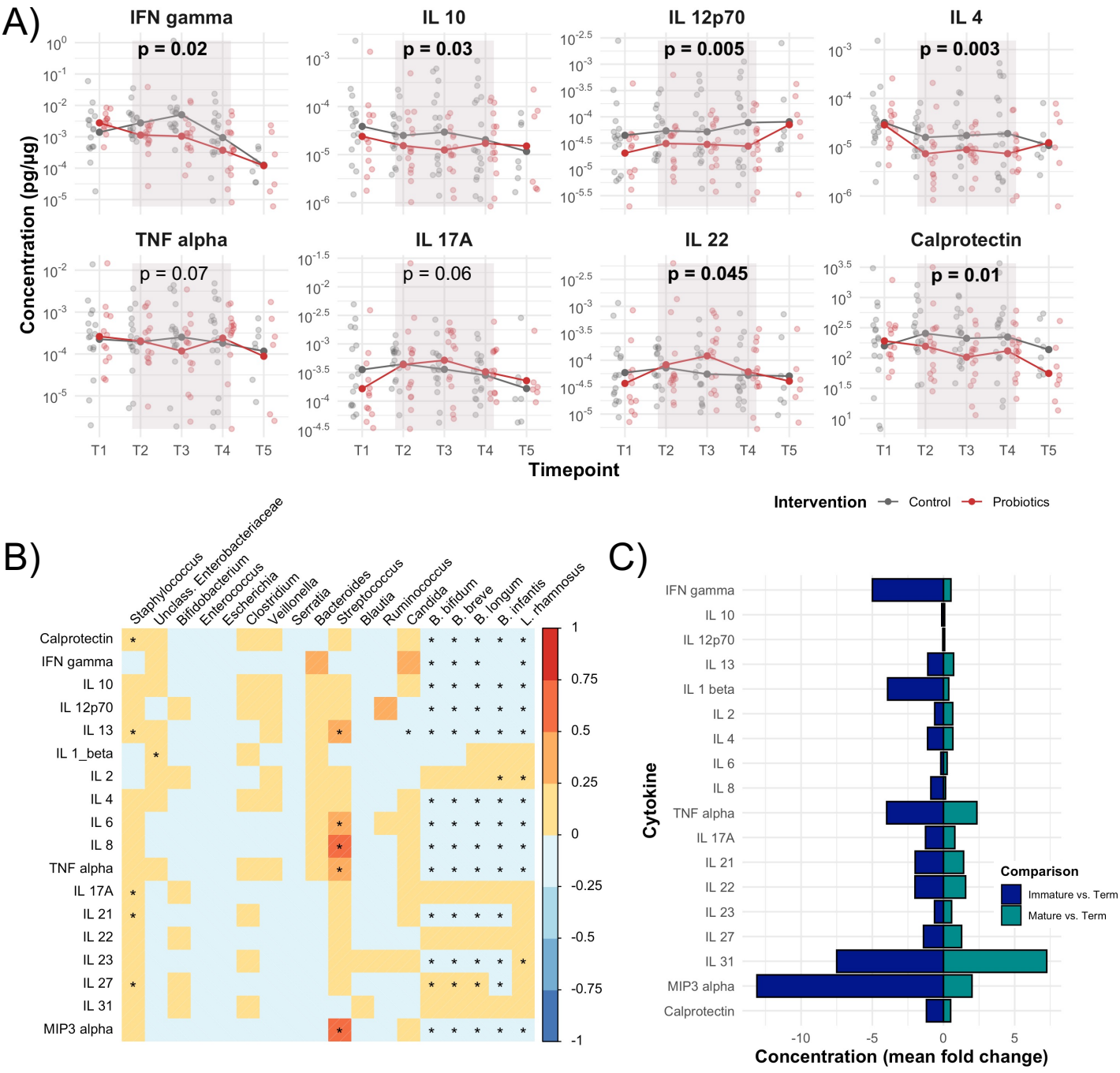


Figure 7. Probiotic-induced microbiome maturation reduced proinflammatory cytokines in stool of extremely premature infants. A) Cytokine concentrations in premature infants according to the intervention. Comparisons were made by generalized estimating equation (Extended Data Table S6). **B)** Correlation of fecal cytokine levels with the 12 most abundant bacterial genera (mean relative abundance > 1%), *Candida*, and probiotic strains log 10 transformed cell numbers. Statistical significance was assessed by adjusting for multiple comparison using Benjamini and Hochberg method. **C)** Cytokine levels by microbiome maturity in relation to term breastfed infants. Mean fold difference in the mature-term vs. immature-term comparisons are shown .

Figure S1

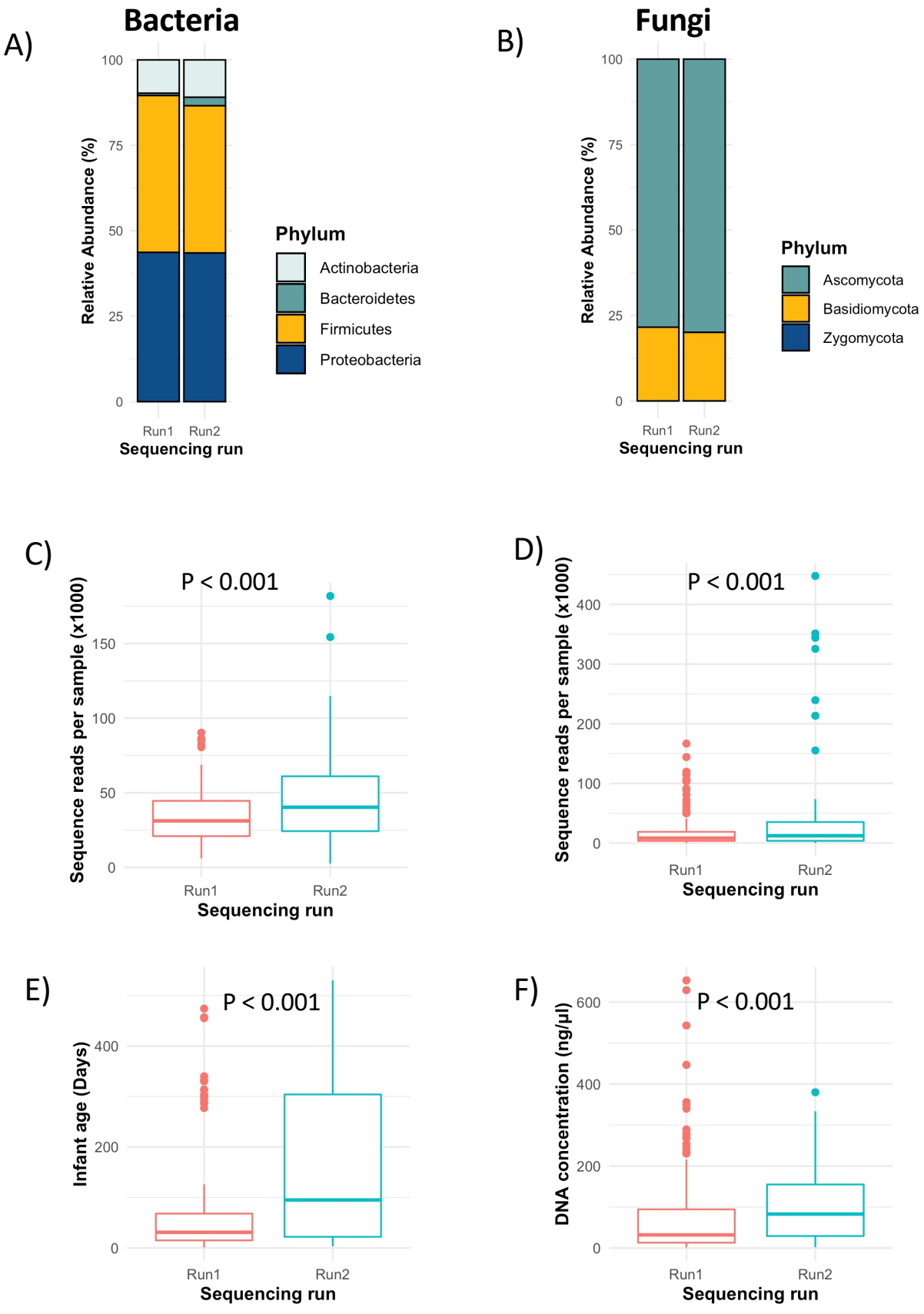

A)
Bacteria
Relative Abundance (%)
100
75
50
25
0
Run1
Run2
Sequencing run
Phylum
Actinobacteria
Bacteroidetes
Firmicutes
Proteobacteria
B)
Fungi
Relative Abundance (%)
Sequencing run
Phylum
Ascomycota
Basidiomycota
Zygomycota
C)
P < 0.001
Sequence reads per sample (x1000)
150
100
50
0
Run1
Run2
Sequencing run
D)
P < 0.001
Sequence reads per sample (x1000)
400
300
200
100
0
Sequencing run
E)
P < 0.001
Infant age (Days)
400
200
0
Sequencing run
F)
P < 0.001
DNA concentration (ng/µl)
600
400
200
0
Sequencing run

Figure S1. Sequencing technical accuracy verification. A-B) Composition of biological controls at phylum level correspond strongly between the two sequencing runs for both bacteria **(A)** and fungi **(B). C-D)** Depth of sequencing is higher in Run 2 in both bacteria 16S rRNA gene sequencing **(C)** and fungi ITS2 sequencing **(D). E)** Run 2 is enriched in older infants and **F)** Run 2 has higher total genomic DNA concentration.

Figure S2

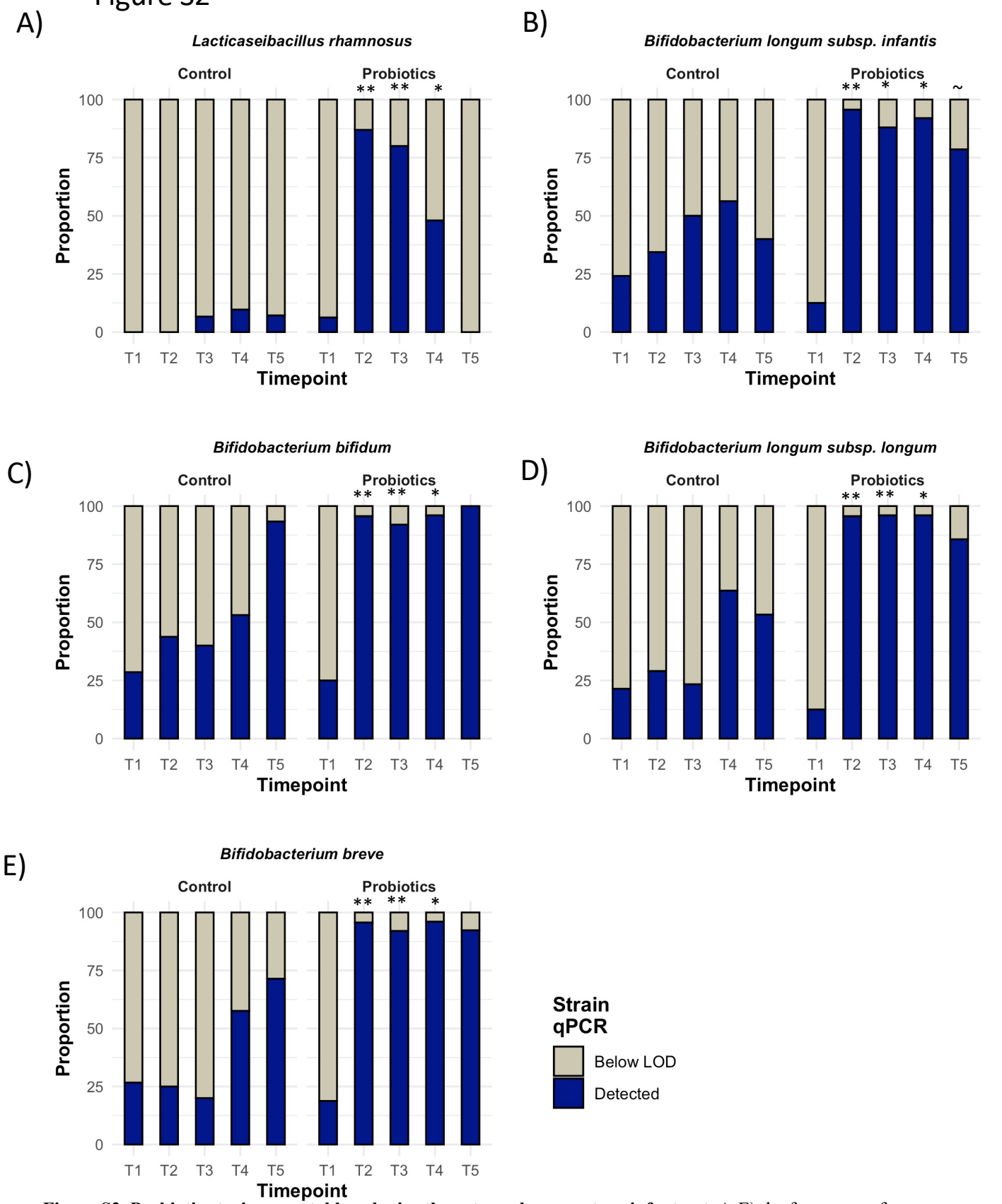
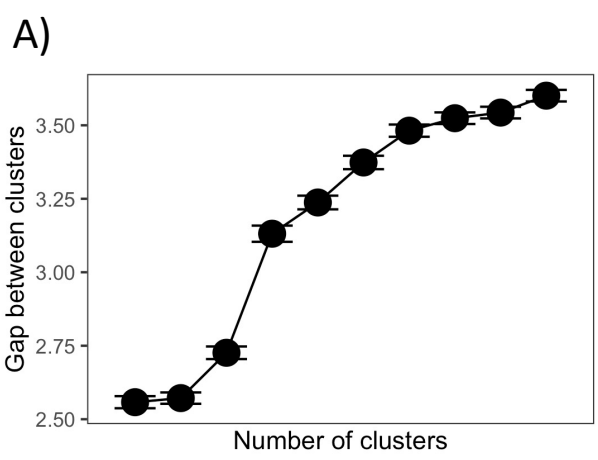
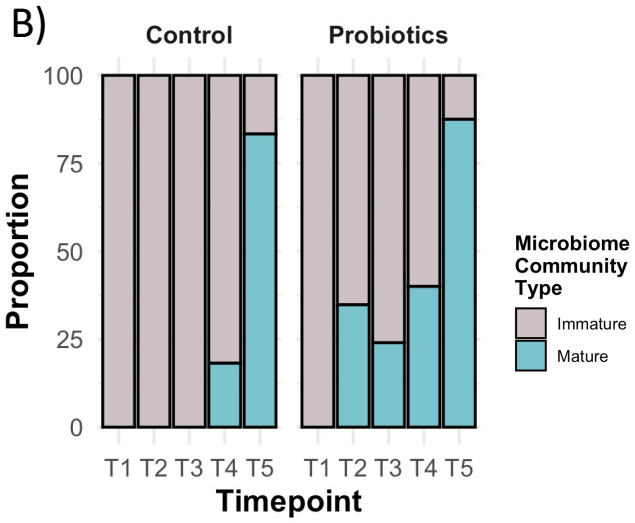
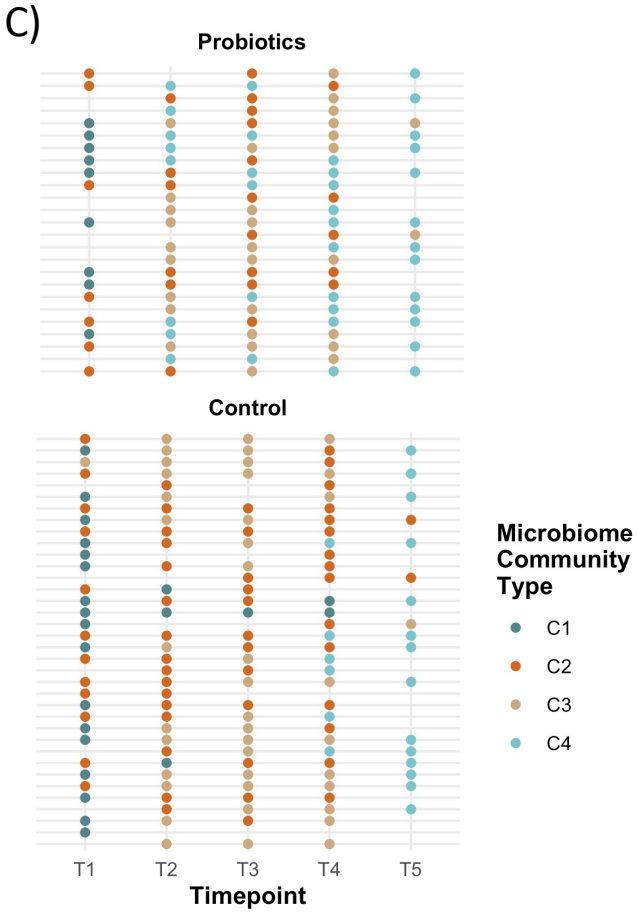


Figure S2. Probiotic strains can stably colonize the extremely premature infant gut. A-E) the frequency of probiotic strains detection at different timepoints are compared between controls and infants who received the probiotic. The frequencies were statistically tested between the intervention groups at each timepoint using $\chi 2$ test. ~ $p < 0.01$, * $p < 0.05$, ** $p < 0.001$.

Figure S3


A)
Gap between clusters
3.50
3.25
3.00
2.75
2.50
Number of clusters


B)
Control
Probiotics
Proportion
100
75
50
25
0
T1 T2 T3 T4 T5
T1 T2 T3 T4 T5
Timepoint
Microbiome Community Type
Immature
Mature


C)
Probiotics
Control
T1
T2
T3
T4
T5
Timepoint
Microbiome Community Type
C1
C2
C3
C4

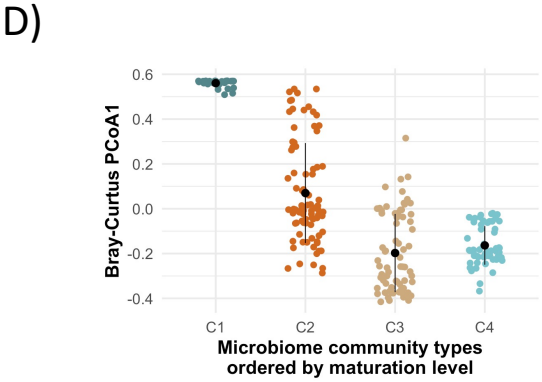

D)
Bray-Curtus PCoA1
0.6
0.4
0.2
0.0
-0.2
-0.4
C1
C2
C3
C4
Microbiome community types ordered by maturation level

Figure S3. Microbiome community type assessment in the bacterial microbiome of preterm infants. A) The optimal number of clusters was identified using Gap statistics. Subsequently, microbiome community types were identified by hierarchical clustering and cutting the tree into 4 clusters**. B)** Infants who received probiotic arrive at the mature community type as early as T2 while the control group shows delays in microbiome maturation. Mature community type was identified based on the frequency of the 4 clusters across the timepoints (see **Figure 2B**). **C)** Individualized microbiome community trajectories are illustrated in rows, with each row representing individual time series per study participant. **D)** Comparison of bacteria beta diversity using Bray-Curtis dissimilarity in community types. Trend analysis was conducted using trendyspliner in SplinectomeR package[70] (See **Figure 2C**).

Figure S4

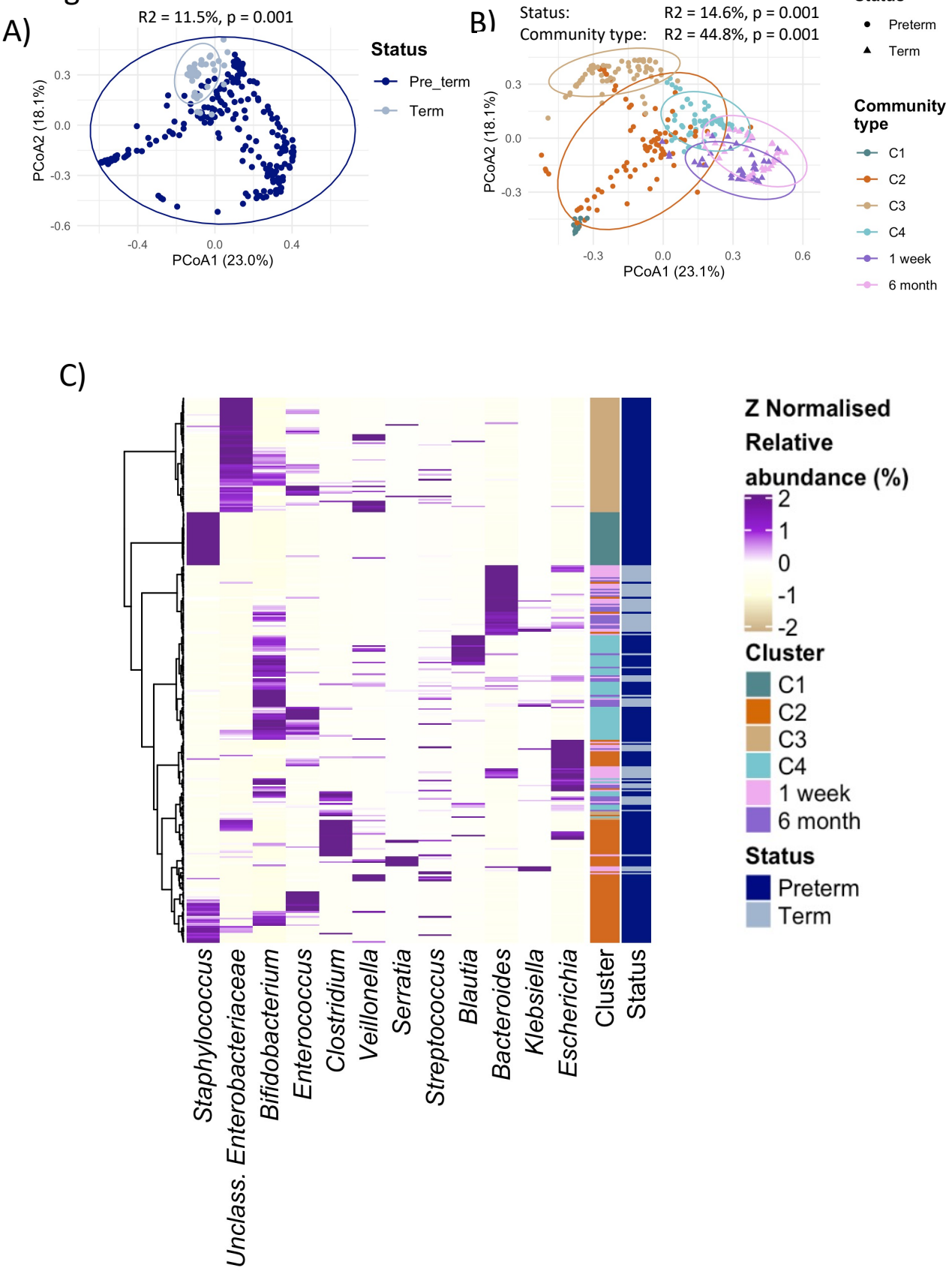

A)
R2 = 11.5%, p = 0.001
PCoA2 (18.1%)
PCoA1 (23.0%)
Status
Pre_term
Term
B)
Status: R2 = 14.6%, p = 0.001
Community type: R2 = 44.8%, p = 0.001
PCoA2 (18.1%)
PCoA1 (23.1%)
Status
Preterm
Term
Community type
C1
C2
C3
C4
1 week
6 month
C)
Z Normalised Relative abundance (%)
2
1
0
-1
-2
Cluster
C1
C2
C3
C4
1 week
6 month
Status
Preterm
Term
Staphylococcus
Unclass. Enterobacteriaceae
Bifidobacterium
Enterococcus
Clostridium
Veillonella
Serratia
Streptococcus
Blautia
Bacteroides
Klebsiella
Escherichia
Cluster
Status

Figure S4. Comparison of the preterm infant microbiome to term breastfed infants.

A) Comparison of the overall composition using Bray-Curtis dissimilarity and visualized as principal coordinate plot (PCoA) reveals that communities are distinct in term vs. preterm infants with the term community having lower heterogeneity. **B).** Comparison of the community types with terms infants using Bray-Curtis dissimilarity and visualized as PCoA plot. The mature C4 community type is closer in composition to the term infants. Association of the infant status and community types with beta diversity was tested using PERMANOVA in panels A and B. **C)** Most abundant taxa at the genus level are visualized in a heatmap comparing community types in preterm infants to the composition of the term infants. Mature C4 community type is clustered with the term infants and is characterized by higher abundance of *Bifidobacterium* spp.

Figure S5

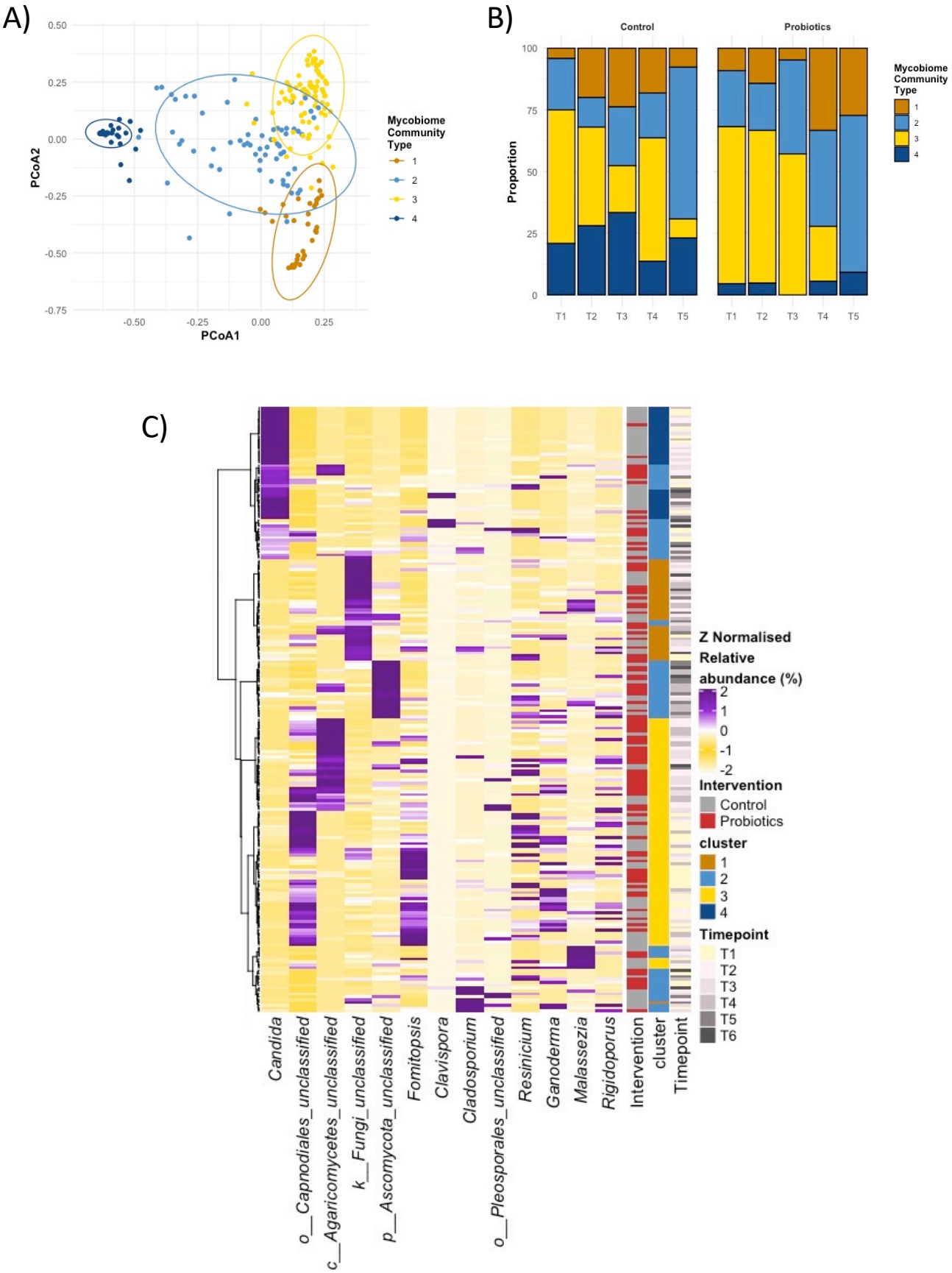

A)
PCoA2
PCoA1
Mycobiome Community Type
1
2
3
4
B)
Control
Probiotics
Proportion
T1
T2
T3
T4
T5
C)
Z Normalised Relative abundance (%)
Intervention
Control
Probiotics
cluster
Timepoint
T6
Candida
o__Capnodiales_unclassified
c__Agaricomycetes_unclassified
k__Fungi_unclassified
p__Ascomycota_unclassified
Fomitopsis
Clavispora
Cladosporium
o__Pleosporales_unclassified
Resinicium
Ganoderma
Malassezia
Rigidoporus

Figure S5. Mycobiome community types are not associated with microbiome maturation trajectory in premature infants. A) Four gut mycobiome community types were identified using hierarchical clustering on Bray-Curtis dissimilarity matrix. **B)** The mycobiome community types do not show strong temporal distribution across timepoint but C4 community type (blue) is reduced throughout the study. **C)** Most abundant taxa at the genus level are visualized in a heatmap according to the intervention, community type (cluster) and timepoints (T1-T5).

Figure S6

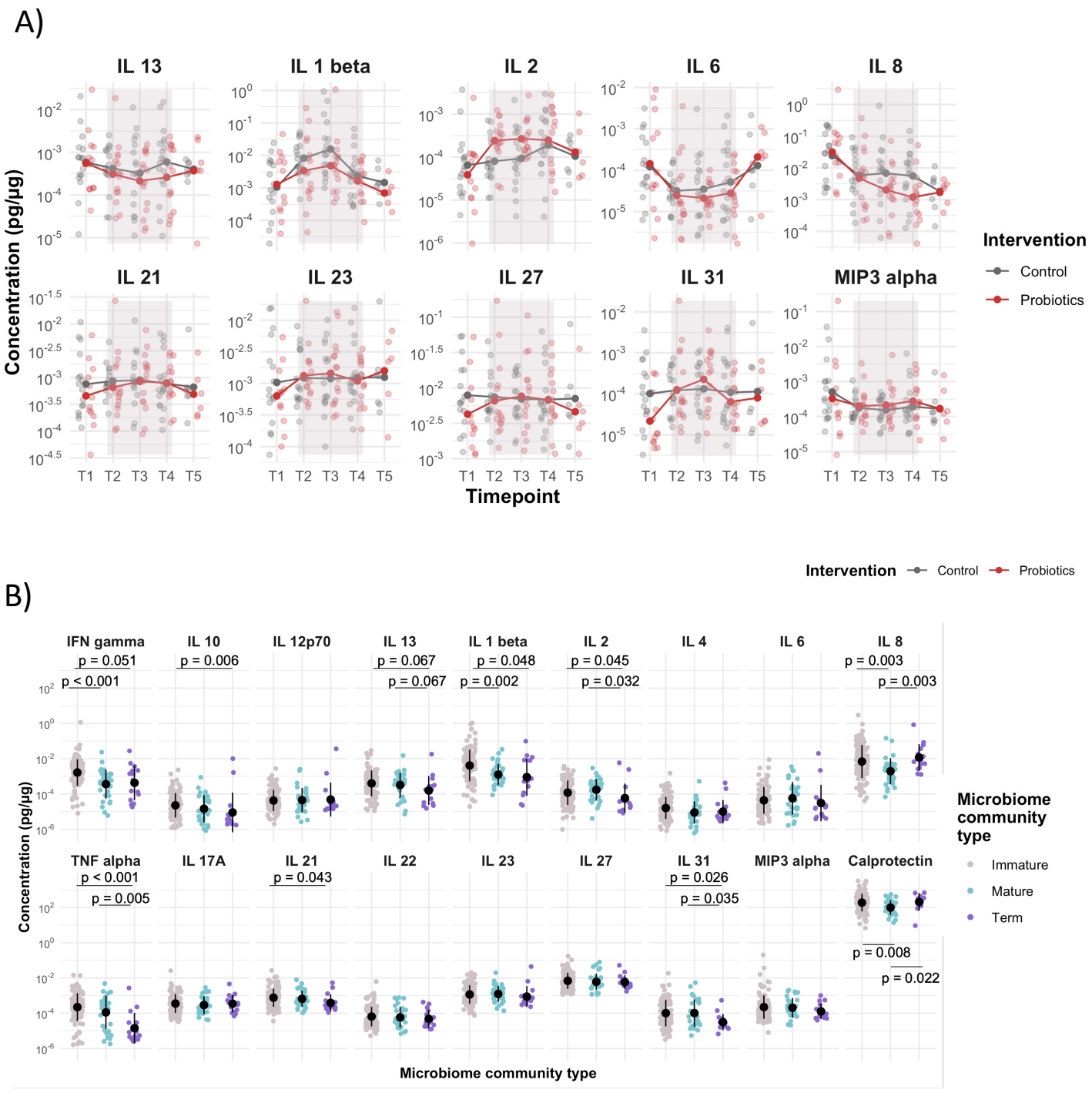


Figure S6. Probiotics reduced proinflammatory cytokines in stool of extremely premature infants. A) Cytokine concentrations in premature infants treated with probiotics (red) and untreated controls (gray) during the randomized clinical trial. Shaded area denotes the hospitalization period after after the probiotic intervention started (Timepoints T2-T4). Comparisons were made by generalized estimating equation **B)** Cytokine concentrations in premature infants according to maturational status. Comparisons were made by pairwise Wilcoxon test.

Probiotic supplementation accelerates gut microbiome maturation in extremely preterm infants (Samara et al.)

Supplementary Tables and Figures:

Table S1. Effect of probiotic use and sample time point on probiotic strain concentration[a] (see Figure 1).

	B. bifidum		***B. breve***		***B. longum***		***B. infantis***		***L. rhamnosus***	
Variables	*Estimates*	*p-value*	*Estimates*	*p-value*	*Estimates*	*p-value*	*Estimates*	*p-value*	*Estimates*	*p-value*
Linear Mixed Model										
Intervention										
Control	Ref.	0.989	Ref.	0.965	Ref.	0.950	Ref.	0.775	Ref.	0.934
Probiotics	0.005		0.016		-0.024		-0.113		0.021	
Timepoint										
T1	Ref.	-	Ref.	-	Ref.	-	Ref.	-	Ref.	-
T2	0.038	0.895	0.066	0.824	0.083	0.789	0.154	0.625	0.003	0.987
T3	0.069	0.815	0.006	0.984	0.154	0.625	0.536	0.096	0.166	0.425
T4	0.847	**0.004**	1.330	**<0.001**	1.071	**<0.001**	0.711	**0.025**	0.151	0.466
T5	1.732	**<0.001**	1.071	**0.005**	0.426	0.268	0.696	0.079	0.013	0.961
Interaction										
Intervention*T2	2.527	**<0.001**	3.106	**<0.001**	3.151	**<0.001**	2.837	**<0.001**	2.606	**<0.001**
Intervention*T3	2.537	**<0.001**	2.893	**<0.001**	2.745	**<0.001**	2.087	**<0.001**	2.100	**<0.001**
Intervention*T4	2.537	**<0.001**	1.827	**<0.001**	2.264	**<0.001**	2.681	**<0.001**	0.676	**0.039**
Intervention*T5	1.250	**<0.001**	1.954	**<0.001**	1.481	**0.012**	0.925	0.125	-0.045	0.907
Post-estimation of linear combination of coefficients										
Cumulative effect of probiotic at:										
T1+T2	2.532	**<0.001**	3.123	**<0.001**	3.127	**<0.001**	2.725	**<0.001**	2.627	**<0.001**
T1+T3	2.541	**<0.001**	2.908	**<0.001**	2.722	**<0.001**	1.975	**<0.001**	2.121	**<0.001**
T1+T4	2.542	**<0.001**	1.843	**<0.001**	2.241	**<0.001**	2.568	**<0.001**	0.697	**0.001**
T1+T5	1.255	**0.005**	1.970	**<0.001**	1.457	**0.001**	0.813	0.083	-0.024	0.936

[a] The analysis was conducted using linear mixed model following log10 transformation of the probiotic cell number. Cell number values below the detection limit (10^3 cells/ml for all probiotic strains) were substituted with limit of detection divided by square root of 2 to account for variance in statistical tests and models.

Table S2. Florababy strain-specific primers

Strain	Primer Name	Primer Sequence	Annealing Temp (˚C)	Amplicon Size (bp)
***Bifidobacterium breve* HA-129**	HA-129_225-F2 HA-129_225-R2	CGACCCTAATGACGTGGAGG CATTTCAGCCAGTACGTGCG	60	195
***Lacticaseibacillus rhamnosus* HA-111**	113A29_293FL 113A29_321RU	ACTCCAAAGAGCATTACCTCCG TGAATATGCCGGATCTAAGTCCA	60	71
***Bifidobacterium bifidum* HA-132**	R71_GB_NC2_F R71_GB_NC2_R	AAGTGTGAGCCGGTGATAGC CAGTACGTCGGCCGTTACAT	60	78
***Bifidobacterium longum subsp. infantis* HA-116**	R33_GB_GE1_F R33_GB_GE1_R	ACGATGCGAGGTGCGATTAT CCCAAGACAAGTCCGCAGAT	60	80
***Bifidobacterium longum subsp. longum* HA-135**	R175_AP_HP10_F R175_AP_HP10_R	GTCGCCACATTTCATCGCAA GAGAGCTTCGATTGGCGAAC	60	99

Table S3. Study participants characteristics

Variables	Probiotics N=26	Control N=31	P value
GA weeks mean (SD)	26 (1)	26 (1)	0.3
BWT mean (SD)	797 (208)	751 (132)	0.324
Sex n (%) Male **Female**	14 (44%) 12 (46%)	11(32%) 21 (68%)	0.4
Multiples n (%)	7 (23%)	10 (29%)	0.77
Mode of delivery n (%) SVD **CS**	8 (31%) 18 (69%)	6 (19%) 25 (81%)	0.37
Birth order n (%) 1st **2nd-3rd** **≥4th**	20 (77%) 4 (15%) 2 (7.7%)	18 (58%) 12 (39%) 1 (3%)	0.13
Chorioamnionitis n (%)	3 (11.5%)	5 (16%)	0.7
PPROM n (%)	9 (35%)	13 (42%)	0.79
Antenatal ABX n (%)	10 (38%)	11(35%)	>0.99
ABX any	23 (89%)	29 (94%)	0.65
EARLY ABX n (%)	19 (73%)	26 (84%)	0.35
Hospital days mean (SD)	106 (26)	105 (28)	0.88
DOT median (25%-75%)	15.5 (8-23)	11.5 (5-17)	0.12

ABX: antibiotics, Early ABX: first 48 hours of life, DOT: days of ABX treatment: sum of duration of each ABX per 100 hospital days. P value was determined by unpaired t test, Mann-Whitney and Fisher exact.

Table S4: Study participant dietary information at ages 1, 2, 4 weeks of life, at 40 weeks PMA and 6 months CA.

Time of Assessment	Participants analyzed Probiotics	Participants analyzed Control	P value
Week 1	**N=26**	**N=31**	
MOM	16 (61%)	18 (58%)	>0.999
DHM	19 (73%)	20 (65%)	0.5737
Week 2	**N=23**	**N=31**	
MOM	19 (83%)	27(87%)	0.7108
DHM	6(26%)	7 (29%)	>0.999
HMF	22 (96%)	26 (84%)	0.2241
Week 4	**N=25**	**N=28**	
MOM	21 (84%)	26 (93%)	0.4042
DHM	3(12%)	4(14%)	>0.9999
HMF	21(84%)	26(93%)	0.4042
PMA 40 weeks	**N=25**	**N=30**	
Exclusive MOM	15 (60%)	23(77%)	0.2447
Mixed (MOM+ Formula)	4(16%)	3(10%)	0.24
Formula	10(40%)	7 (23%)	0.2447
6 mo CA	**N=24**	**N=30**	
MOM 6mo CA	5(21%)	7(23%)	>0.9999
Formula	21(88%)	27(93%)	>0.9999

MOM: mother's own milk, DHM: donor human milk, HMF: human milk fortifier, Formula: any type of artificial formula milk.

Table S5. Differential metabolites in premature infant stool between immature and mature community types

Compounds	Fold change (FC)	Log2 FC	Adjusted p (FDR)	-log10(p)
Leucine	3.8787	1.9556	7.7103e-05	4.1129
N-Acetyl-DL-glutamic acid	0.41038	-1.285	0.00049245	3.3076
L-phenylalanine	2.3824	1.2524	0.00049245	3.3076
N-Acetyl-L-aspartic acid	0.34299	-1.5438	0.0032034	2.4944
Inosine	5.704	2.512	0.0035457	2.4503
L-valine	2.3677	1.2435	0.0035574	2.4489
Oleate	10.025	3.3255	0.01729	1.7622
Orotate	0.2987	-1.7432	0.017844	1.7485
Arachidic acid	6.3829	2.6742	0.018335	1.7367
L-cysteine	3.3505	1.7444	0.018335	1.7367
Palmitoleic acid	10.038	3.3274	0.020116	1.6965
Cholate	0.044846	-4.4789	0.038278	1.417
LL-2,6 diaminoheptanedioate	0.34884	-1.5194	0.043007	1.3665
Taurine	0.23445	-2.0927	0.043544	1.3611

Features selected by fold change threshold (2) and false discovery rate (0.05).

Table S6. Effect of probiotic and sample timepoint on stool cytokine concentrations.

The analysis was conducted using generalized estimating equation (GEE). The optimum GEE model for each cytokine was selected based on the cytokine distribution and the model performance with different correlation structures: independence, exchangeable, autoregressive 1, or unstructured. The family of the GEE model was set as gaussian or gamma for normal or positively skewed cytokine distribution, respectively.

	Intervention [probiotics]		Timepoint [T3]		Time point [T4]	
Cytokine (pg/mg total protein)	Estimate	p-value	Estimate	p-value	Estimate	p-value
Calprotectin	-0.57	**0.010**	-0.28	0.219	-0.13	0.592
IFN gamma	-1.44	**0.019**	1.43	0.129	-0.83	0.134
IL-10	-0.84	**0.031**	0.01	0.975	0.25	0.623
IL-12p70	-0.98	**0.005**	-0.07	0.860	0.54	0.256
IL-13	-0.54	0.250	-0.36	0.557	0.07	0.909
IL-1 beta	-0.26	0.638	1.29	**0.019**	-1.44	**0.006**
IL-2	0.46	0.133	0.34	0.233	0.84	**0.001**
IL-4	-1.13	**0.003**	-0.03	0.941	0.44	0.388
IL-6	-0.80	0.193	-0.74	0.429	-0.41	0.613
IL-8	-0.93	0.100	-0.21	0.796	-1.14	0.075
TNF alpha	-0.54	0.069	0.21	0.512	0.21	0.519
IL-17A	0.67	0.064	-0.48	0.330	-0.89	0.061
IL-21	0.14	0.711	-0.05	0.922	-0.19	0.685
IL-22_	0.86	**0.045**	-0.47	0.351	-0.71	0.175
IL-23	-0.02	0.933	0.12	0.728	-0.10	0.797
IL-27	0.34	0.245	-0.16	0.669	-0.17	0.667
IL-31	0.49	0.398	-0.49	0.408	-0.46	0.513
MIP3 alpha	0.56	0.105	0.09	0.831	0.76	0.102