Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

SCIENZE BIOTECNOLOGICHE E FARMACEUTICHE

Ciclo XXX

Settore Concorsuale: 03/D1

Settore Scientifico Disciplinare: CHIM/11

Dynamics of the infant microbiomes onset: exploring the gut and the oral microbiota in full-term and pre-term infants in the frame of mother's milk microbial ecosystem

Presentata da: **Sara Quercia**

Coordinatore Dottorato — **Prof. Santi Mario Spampinato**

Supervisore — **Prof. Patrizia Brigidi**

Esame finale anno 2018

Table of contents

Abstract

The gut microbiota assembly during the very first days of life plays a pivotal role in the education of the immune system and in the building of a healthy status later in life. Besides the mode of delivery, also the feeding type and the gestational age have an impact on the gut microbiota composition. For this reason, by means of next-generation sequencing of the 16S rRNA gene on Illumina MiSeq, we characterized and compared the intestinal bacterial community in 2 cohorts of infants, 1 constituted by 36 healthy breast-fed infants born full-term and the other constituted by 21 infants born moderate to late pre-term (32 to 37 weeks) receiving different types of feeding (mother breast milk, human breast milk from donor and formula). The first cohort was sampled at 20th day of life, whilst the second one was sampled longitudinally from birth to 30th day of life. In addition, also the infant's saliva and the mother's milk were sampled and sequenced.

The characterization of the 3 ecosystems in full-term infants led to the hypothesis that the mother milk, together with the microorganisms that reside in the baby's mouth, may act as seeding community and may participate to infant gut microbiota assembly. On the other hand, in the moderately-to late pre-term cohort, the extreme diversity of the infant's clinical history provokes a tremendous inter-individual variability. Nevertheless, milk and saliva microbiological structures resembled the ones of full-term cohort. The gut microbiota instead presented a very different composition and it is plausible that its establishment is strongly influenced by the infant's clinical history and environmental bacteria than the mutual relationship with the mother.

Introduction

Almost 150 years have passed since Il'ja Il'ič Mečnikov theorized that health could be enhanced by manipulating the gut microbiota (GM). Where do we stand today? Many studies have been carried out to investigate the characteristics the intestinal community and of the other microbial ecosystems that inhabit human body. Our organism contains up to 27 sites where a microbial community can be found (Costello *et al.*, 2009). The gastrointestinal tract is fully colonized by microorganisms, starting from the mouth down to the colon. The oral cavity represents the first access to the gastrointestinal tract and it communicates perpetually with the external environment. In the mouth, many different ecosystems can be found: salivary, gingival, lingual and mucosal. The salivary microbiota (SM) of a healthy adult is scarcely biodiverse and it is mainly constituted by genus *Streptococcus* (species *S. salivarius* and *S. mitis*), but also the genera *Neisseria*, *Rothia* and *Prevotella* are present (Yun-ji Kim *et al.*, 2016; Zaura *et al.*, 2014). Yet the most dense and biodiverse of these communities present in the GIT resides in the colon, where it reaches the concentration of 10^{12} CFU/g of luminal content. When the GM is described at a compositional level, only a limited number of phyla is found: Firmicutes, Actinobacteria, Bacteroidetes, Proteobacteria, and Verrucomicrobia, with Firmicutes and Bacteroidetes accounting for up to 90% of the ecosystem (Costello *et al.*, 2009). On the contrary, if we look at a lower phylogenetic level, the biodiversity explodes, reaching more than 1000 species (Qin *et al.*, 2010). The microbial community that resides in our gut is very unique, describing a personal fingerprint. The GM provides the host with metabolic functions, such as the digestion of complex polysaccharides, production of vitamins, cofactors and other secondary metabolites (Bäckhed *et al.*, 2004). There is a wide range of metabolites produced that depend on the macronutrient ingested. The endpoint of polysaccharide fermentation is mainly represented by short-chain fatty acids (SCFAs), namely acetic, propionic and butyric. They participate in the pathway that regulates appetite and play a role in host nutrition and energy homeostasis, controlling energy production and storage (Russel *et al.*, 2013). On the other hand, when amino acids are metabolized by the GM, we obtain indolic and phenolic compounds, together with methylamines, which are linked to obesity, type 2 diabetes and hepatic steatosis (Lin *et al.*, 2017).

Many researches have also demonstrated how crucial the GM is in educating the immune system starting from the very beginning of life, by modulating the generation of an equilibrium between anti- and pro-inflammatory response when exposed to bacteria (Berrington *et al.*, 2013). An impaired GM leads to a defective communication with the immune system, causing persisting effect on host physiology later in life even after GM resembling (Arrieta *et al.*, 2014).

The gut colonization was thought to start during delivery, but recently Aagard and colleagues demonstrated that microbial DNA belonging to the species *Escherichia coli*, *Bacteroides* spp., *Neisseria lactamica*, *Staphylococcus epidermidis* and *Propionibacterium acnes* is found in the placenta (Aagard *et al.*, 2014). Nevertheless, the mode of delivery, together with gestational age at birth and type of nutrition, play a pivotal role in the bacterial establishment at early stage. The most favourable condition for the infant health status is being born at term with vaginal delivery and being fed exclusively with maternal breast milk. However, statistics says that in Italy in 2015, 34.1% of the delivery were C-sections (Italian Health Ministry Data). Vaginal-delivered infant GM is different from the C-section one (Dominguez-Bello *et al.*, 2010; Gritz and Bhandari, 2015): while vaginally-born infants have a gut microbial composition enriched in *Bacteroides, Bifidobacterium, Parabacteroides, Escherichia* and *Shigella* (Bäckhed *et al.*, 2015) and they present also bacteria deriving from maternal vaginal microbial community, such as *Lactobacillus* and *Prevotella* species (Gritz and Bhandari, 2015), infants born via C-section are instead colonized by epidermal and environmental species, namely *Clostridium*, *Staphylococcus*, *Propionobacterium*, and *Corynebacterium* and, when they are compared to the vaginally-born, they have lower levels of anaerobes, in particular *Bacteroides* and *Bifidobacterium* (Brugman *et al.*, 2015). These differences though smooth over with the introduction of milk. Newborns can be fed with maternal breast milk, human breast milk from donor and formula. Human breast milk is thought to be the most beneficial for the baby's health and also for the GM establishment (Agostoni *et al.*, 2009). It contains a mixture of nutrients and molecules, such as immunoglobulins (IgA), carbohydrates, fatty acids and lactoferrin (Jain and Walker, 2014). Among the glucidic fraction, the human milk oligosaccharides (HMOs) represent one of the most investigated, because of its capacity to selectively stimulate the growth of specific bacteria in the infant gut. Human milk has not only a prebiotic function, but it has been suggested that could act also as microbes source, being enriched in *Streptococcaceae*, *Bifidobacteriaceae* and *Staphylococcaceae* (Biagi *et al.*, 2017). The origin of these bacteria has not been elucidated yet. An enteromammary pathway has been suggested by Perez and colleagues (2006) and they support the idea that fragments of DNA, antigens and microbial protein may be transported via blood by dendritic cells to the mammary gland and be secreted in the milk. Throughout the milk, the mother starts to educate the immature immune system of the baby to the late formation of a microbial community. Another theory explains the presence of bacteria as a contamination coming from the infant's mouth, that inoculate milk duct during suction (Biagi *et al.*, 2017). Infants receiving nourishment from the mother display a GM structure dominated by species of *Bifidobacterium* and, when compared to formula-fed infants, show lower relative abundance of *Enterobacteriaceae*. Formula-fed ones also show the presence of *Escherichia coli, Clostridium difficile, Bacteroides*,

Prevotella, and *Lactobacillus* (Jost *et al.*, 2012). C-section is a more common practice in premature birth, when it is essential to give birth to the foetus due to diseases affecting either the mother or the baby, as well as uterine infections or breakage of the amniotic sac. A baby born premature is usually described according to the gestational age, that is the weeks dating from the first day of the mother's last menstrual period (NIH, US library of medicine). The World Health Organization has drawn up a classification to define newborns according to gestational age:

- extremely pre-term (<28 weeks);
- very pre-term (28 to <32 weeks);
- moderate to late pre-term (32 to <37 weeks);
- full-term (37 to 40 weeks).

It is estimated that every year in Italy around 40.000 babies are born pre-term (Italian Society of Neonatology). Pre-term infants display a different gut colonization when compared to full term ones (Arboleya *et al.*, 2012) and it is influenced not only by the mode of delivery (mainly C-section) and nutrition (mainly with a mixture of human breast milk and formula), but also by an impaired feeding process, due to the delay in the full establishment of coordinated latch, suckling, swallowing and breathing (Mizuno *et al.*, 2003), incomplete oesophageal peristalsis (Staiano *et al.*, 2007) and altered gastric emptying (Riezzo *et al.*, 2000). These newborns are exposed to many complications, such as necrotizing enterocolitis (NEC) and sepsis (Gregory *et al.*, 2016). NEC is characterized by intestinal inflammation that can lead to tissue necrosis and sepsis. Its onset appears to be multifactorial: gut immaturity, intestinal damage or injury, enteral feeding and impaired bacterial colonization play a role, although the exact pathogenesis remains unidentified. The role of GM establishment in NEC has been investigated by a meta-analysis carried out by Pammi and colleagues (2017) and they highlighted that the onset of the disease is predated by an increase of Proteobacteria and a concomitant decrease of Firmicutes and Bacteroidetes. Moreover, maternal breast milk administration is considered protective against NEC onset (Meinzen-Derr *et al.*, 2009).

Although many researches focus on GM features in very and extremely pre-term infants, moderately to late pre-term is considered a neglected category, because poorly investigated. Even if they have a much lower risk of medical complications than more premature infants, they still experience higher rates of infant morbidity and mortality, as well as higher risks of childhood disabilities (Shapiro-Mendoza *et al.*, 2012) when compared to full term babies.

For these reasons it is important to focus the research on this group, describing their gut ecosystem from birth until weaning, in the frame of milk microbial composition. Finally, in order to understand how much these babies differ from the normal healthy situation, it is crucial to compare them with full term vaginally delivered babies, who are exclusively breastfed. This approach allows the

identification of the bacterial genera considered "health-promoting" that are lacking in the GM of pre-term infants, paving the way to the development of an individual probiotic intervention to restore the ecosystem.

Project outline

It's becoming always clearer that disturbances that strike the GM in newborns affect health status later in life. Many factors can interfere with a correct GM establishment, such as gestational age, mode of delivery and type of feeding. According to gestational age, we can divide the newborns in 4 categories (WHO definition):

- extremely pre-term (<28 weeks)
- very pre-term (28 to <32 weeks)
- moderate to late pre-term (32 to <37 weeks).
- full-term (37 to 40 weeks)

Being born full-term with vaginal delivery and being fed exclusively with mother breast milk is considered the most desirable condition for health and for a correct GM establishment. For this reason, the first part of this research work was addressed to the characterization of the gut microbial community of 36 healthy, vaginally delivered and full-term babies, who were exclusively breastfed at 20th day of life. In addition, also saliva microbiota from the baby and breast milk from the mother were analysed, in order to understand the relationship between the infant GM and the mother milk microbial ecosystem, in the frame of the mouth community, which acts as mandatory connection.
Consequently, our focus moved to infants who are born between 32 to 37 weeks (moderate to late pre-term). This group is considered a neglected one, because yet poorly investigated. According to the OMS report on pre-term births drawn up in 2012, the total number of premature infants was of 131,296,785 and the moderate to late pre-term category accounted for the 84.3%. These newborns do not have the typical clinical progress of full term babies, having high probability of developing many complications, in particular necrotizing enterocolitis (NEC) and sepsis. Nevertheless, their complication onset rate is smaller when compared to the very and extremely pre-term ones. GM assembly appears to play a pivotal role in the development of such complications and for this reason we decided to characterize the gut microbial community establishment in a cohort of 21 neonates sampled longitudinally. From these babies, in addition to faecal samples, also a saliva swab was taken, together with an aliquot of mother's milk (when present) at different timepoints, from birth until weaning, collecting a total number of 348 samples. The microbial community of all these samples was sequenced using Illumina platform, then characterized and described from a phylogenetic point of view; moreover, the impact of breastfeeding was examined. Finally, the comparison between full term and pre-term babies at 20 days of life was performed in order to decipher the differences between these 2 conditions.

Experimental procedures

Total bacterial DNA extraction from complex matrix

Total bacterial DNA was extracted from feces using the DNeasy Blood and Tissue Kit (QIAGEN, Hilden, Germany) with a modified protocol (Yu and Morrison, 2004). Briefly, 250 mg of stool samples were resuspended in 1 mL of lysis buffer (500 mM NaCl, 50 mM Tris–HCl pH 8, 50 mM EDTA and 4% SDS) and treated with 3 beads-beating steps in FastPrep instrument (MP Biomedicals, Irvine, CA) at 5.5 movements per sec for 1 min and kept in ice among treatments. Samples were centrifuged at full speed for 5 min at 4°C, then 260 µl of 10M ammonium acetate were added and the samples incubated for 5 min in ice. After 10 min of centrifugation at full speed at 4°C the supernatants were collected and 1 volume of isopropanol added. Samples were mixed and incubated in ice for 30 min. DNA was collected by 15 min of centrifugation at full speed at 4°C and the pellet washed with 70% ethanol. The pellet was then resuspended in 100 µl of TE buffer and RNA and proteins removed treating the samples respectively with 2 µl of DNase-free RNase (10 mg/ml) for 15 min at 37°C and 15 µl of proteinase K at 70°C for 10 min. DNA was further purified using QIAamp Mini Spin columns (QIAGEN) following the manufacturer's instructions. For milk samples, the same protocol described above for fecal samples was applied, preceded by the centrifugation of 2 ml of sample at full speed for 10 min at 4°C. For DNA extraction from oral swabs, the cotton swab was suspended in 500 µl of PBS, vortexed for 1min and sonicated for 2 min. These 2 steps were repeated twice, then 2 cycles of bead-beating with FastPrep at 5.5 movements per sec for 1min, with 200 mg of glass beads, were applied. Cotton residues were removed and the debris pelleted by centrifugation at 9000g for 5 min. The supernatant was discarded and the pellet resuspended in 180 µl of enzymatic lysis buffer (QIAGEN). Samples were then treated according to the DNeasy Blood&Tissue kit (QIAGEN) instructions, following the protocol for Gram positive bacteria. Extracted DNAs were quantified using the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE).

16 rRNA gene amplification and sequencing

For each sample, the V3-V4 region of the 16S rRNA gene was PCR amplified in 25 µl final volume containing 5 µl of microbial DNA (diluted to 5 ng/µl for fecal samples, undiluted for milk and oral swab), 2X KAPA HiFi HotStart ReadyMix (KAPA Biosystems, Resnova, Rome, Italy), and 200nM of S-D-Bact-0341-b-S-17/S-D-Bact-0785-a-A-21 primers (Klindworth *et al.*, 2013) carrying Illumina

overhang adapter sequences. PCR conditions set up as follows: initial denaturation at 95°C for 3 minutes, 25 cycles of denaturation at 95°C for 30 seconds, annealing at 55°C for 30 seconds, and extension at 72°C for 30 seconds, and a final extension step at 72°C for 5 minutes. PCR amplicons were purified with a magnetic bead-based clean-up system (Agencourt AMPure XP; Beckman Coulter, Brea, CA). Indexed libraries were prepared by limited-cycle PCR using Nextera technology and further cleaned up with AMPure XP magnetic beads (Beckman Coulter). Libraries were pooled at equimolar concentrations (4nM), denatured and diluted to 6 pmol/L before loading onto the MiSeq flow cell. Sequencing on Illumina MiSeq platform was performed by using a 2×300 bp paired end protocol, according to the manufacturer's instructions (Illumina, San Diego, CA).

Bioinformatics

Data analysis is performed using a pipeline combining PANDAseq (paired-end assembler for Illumina sequences) (Masella *et al.*, 2012) and QIIME (Quantitative Insights Into Microbial Ecology) (Caporaso *et al.*, 2010). High-quality reads are filtered and then clustered into operational taxonomic units (OTUs) at a 0.97 similarity threshold using UCLUST (Edgar, 2010). Taxonomy is assigned using the RDP (Ribosomal Database Project) classifier against Greengenes database (May 2013 release) and the chimera filtering is performed by discarding all singleton OTUs. Alpha diversity is measured using the Chao1, observed species and Shannon index metrics. Beta diversity was estimated by computing Bray-Curtis distances. The distance matrix obtained was used for principal coordinates analysis (PCoA) and plotted using the rgl and vegan packages of R.

Part A: The bacterial ecosystem of mother's milk and infant's mouth and gut

Biagi Elena, Quercia Sara, Aceti Arianna, Beghetti Isadora, Rampelli Simone, Turroni Silvia, Faldella Giacomo, Candela Marco, Brigidi Patrizia, Corvaglia Luigi

Published on Frontiers in microbiology in 2017

Introduction

The microbiota of individuals with whom a human being has direct and frequent contacts contributes in shaping its microbial communities (Song *et al.*, 2013; Stahringer *et al.*, 2012). This is even more true in the case of breastfed infants and their mothers, where the microbial ecosystems of the latter are the most relevant sources of colonizing microbes for the former (Arrieta *et al.*, 2014). The progressive building of the infants' microbiota, especially for what concerns the gut ecosystem, is a crucial proceeding for educating their immune system to the delicate balance between tolerance and reactivity that is needed to maintain health throughout the entire human life (Arrieta *et al.*, 2015; Honda and Littman, 2016; Lynch and Pedersen, 2016). Consequently, the understanding of the colonization dynamics of the infant's gut microbiota is not only fascinating from the ecological point of view, but also incredibly relevant for clinical immunology (Arrieta *et al.*, 2015; Honda and Littman, 2016).

The infant's gut microbiota is a highly dynamic community that is progressively and continuously shaped during the first days of life, with nutrition (breast vs. formula feeding) being among the most relevant drivers for its composition (Gritz and Bhandari, 2015). With its estimated 3 log CFU/ml of bacterial concentration (Jost *et al.*, 2014), human breast milk is listed among the first sources of microbes for the infant's gut ecosystem, together with the mother's skin, mouth and vaginal tract, in case of vaginal delivery (Mueller *et al.*, 2015). Research struggles to give a conclusive demonstration for the origin of the bacteria recovered in human milk: even if a controversial "bacterial entero-mammary pathway" has been proposed (Rodríguez, 2014), contamination by the surrounding skin microbiota and other environmental sources might also occur. Indeed, facultative anaerobic or prevalently aerobic species are the major colonizers of the human milk ecosystem: *Streptococcus* and *Staphylococcus* are the most frequently isolated and abundant bacterial groups in milk samples, together with skin-derived or environmental contaminants (i.e. *Propionibacterium* and genera of the Enterobacteriaceae family) (Fitzstevens *et al.*, 2016). However, well-known intestinal probiotic bacteria (i.e. *Bifidobacterium* and *Lactobacillus*) are often retrieved by both molecular and

cultivation-based studies (Fitzstevens *et al.*, 2016). Next generation sequencing also allowed the detection of obligate anaerobic, gut-associated genera, such as *Bacteroides*, *Blautia*, *Dorea*, and *Faecalibacterium* (Jost *et al.*, 2013); if alive, these bacteria could act as pioneers in the infant gut for the construction of the adult gut microbiota, which will begin to settle down at weaning (Rodríguez, 2014).

In this scenario of microbial exchange between mother and child, the baby's mouth is unavoidably involved, being the obligate transition point for the milk to reach the gastrointestinal tract. The oral microbiota is a well-characterized portion of the human microbiome. It is usually dominated by *Streptococcus* and *Staphylococcus* in healthy, breastfed term infants; aerobic or facultative anaerobic bacterial taxa, such as *Gemella*, *Actinomyces*, and *Veillonella*, act as later and minor colonizers (Sampaio-Maya and Monteiro-Silva, 2013). The mouth is a particularly exposed ecosystem, anatomically open to the external environment and continuously in contact with air, food, and water. For these reasons, this ecosystem needs to cope with chemical, physical and mechanical fluctuations. The mouth of healthy individuals is not routinely found to be colonized by non-oral microorganisms, possibly because exogenous bacteria lack in specific adhesins and receptors that would enable them to bind to oral surfaces, or are excluded by immune mechanisms (Wade, 2013). On the contrary, evidences of seeding of the baby's gut by the oral microbiome have been provided (Costello *et al.*, 2013; Ding and Schloss, 2014).

In this frame, it is crucial to include the bacteria inhabiting the infant's oral cavity in the complex mechanism of bacterial transfer between the mother's milk microbiota and the infant's gut ecosystem. Indeed, the oral ecosystem might contribute in seeding the gut both directly, through deglutition, and indirectly, by contaminating the mother's milk ducts, during suction.

In an attempt to decipher the relationship between the mother's milk ecosystem and the infant's microbiome, we analyzed, to our knowledge for the first time, the microbial composition of oral, gut and milk ecosystems in a small, yet very homogeneous, cohort of 36 healthy mother-infant pairs. By limiting the influence of confounding variables (e.g. delivery mode, gestational age), our findings shed some light on the relevance of bacterial sharing between these ecosystems.

Subjects recruitment

The Nursery of S. Orsola-Malpighi Hospital in Bologna, Italy, recruited mother-infant pairs meeting the following criteria: (i) vaginal delivery at term (≥37 weeks gestation), (ii) exclusive breastfeeding during the sampling period, (iii) no antibiotic/probiotic exposure of either the mother or the infant

during pregnancy, intrapartum or postnatally. Infants who had or developed clinical conditions that required hospitalization were excluded.

Written informed consent was obtained, in accordance with the Declaration of Helsinki, from each mother before the mother-infant pair was discharged from the nursery (48-72 hours after delivery). Follow-up visits at 20 days of life were scheduled in order to obtain a neonatal fecal sample, two neonatal oral swabs (before and after breastfeeding), and a fresh mother milk sample. Feces were collected from diapers using a standard sterile collection tube. Milk samples were collected with the aid of a breast pump into sterile plastic tubes; prior to collection, mothers were asked to wash the nipple and mammary areola with soap and water. Oral samples were obtained by gently swabbing a sterile cotton-tipped applicator on the inside of the infant's cheek. Samples were immediately delivered to the laboratory using cold packs, then split into aliquots ready for DNA extraction and frozen within few hours from collection. Samples were thawed in batches for processing. All samples were processed within 4 months of receipt. Demographic and clinical data were recorded in a specific case report form. The study was approved by the ethics committee of the S. Orsola-Malpighi Hospital in Bologna (study protocol 53/2014/U/Tess). Methods were carried out in accordance with the approved guidelines.

Experimental procedure and statistics

Samples processing were conducted as described in "Experimental procedure". Statistics was performed using R software (https://www.r-project.org/) and the libraries vegan and made4. Weighted and unweighted UniFrac distances were used for Principal Coordinates Analyses (PCoA), and the significance of separation was tested by permutational multivariate analysis of variance using the function "adonis" of the vegan package, after testing for homogeneity of dispersion using the function "betadisper". Wilcoxon test was used to assess significant differences between two groups of samples; adaptations for paired samples were used when necessary. Kruskal-Wallis test was used for multiple comparisons, followed by Tukey post-hoc test when appropriate. P values were corrected for multiple comparisons using the Benjamini-Hochberg method. $P<0.05$ was considered as statistically significant. Correlation between datasets was tested by using the Kendall method.

Results and discussion

Thirty-six mother-infant pairs were included in the study. All infants were vaginally delivered and exclusively breastfed. Neither infants nor mothers had received any antibiotics or probiotics until the sampling date. One-hundred-forty-three samples were collected 20 days after delivery: 36 mother's milk samples, 36 infant's fecal samples, and 71 infants' oral swabs (35 pairs of pre and post breastfeeding samples plus one unpaired pre-breastfeeding sample).

The extracted bacterial DNA was phylogenetically characterized by 16S rRNA gene (V3-V4 region) Illumina sequencing. A total of 1,475,619 high-quality reads was obtained with a mean of 10,319±3,364 reads per sample. Rarefaction curves obtained with Shannon and Chao1 metrics approximated the saturation level after 3,000 reads. Reads were clustered in 7,524 operational taxonomic units (OTUs) at 97% of identity. OTU table and taxa summary tables at family and genus level are available as Supplementary Tables.

	Mother-infant pairs (%) in which the same OTU was present in ≥2 ecosystems					
OTU ID	**M-Opre-F**	**M-Opost-F**	**M-Opre**	**M-Opost**	**M-F**	**Assigned Taxonomy (within Bacteria kingdom)**
24152	23	20	31	20	46	p_Actinobacteria; c_Actinobacteria; o_Bifidobacteriales; f_Bifidobacteriaceae; g_Bifidobacterium; s_breve
24615	63	63	97	97	63	p_Firmicutes; c_Bacilli; o_Lactobacillales; f_Streptococcaceae; g_Streptococcus; s_infantis
36228	3	3	51	51	3	p_Firmicutes; c_Bacilli; o_Gemellales; f_Gemellaceae; g_; s_
36370	23	37	34	60	60	p_Proteobacteria; c_Gammaproteobacteria; o_Enterobacteriales; f_Enterobacteriaceae; g_; s_
41720	26	29	29	34	74	p_Actinobacteria; c_Actinobacteria; o_Bifidobacteriales; f_Bifidobacteriaceae; g_Bifidobacterium; s_longum
90869	0	0	3	0	51	p_Actinobacteria; c_Actinobacteria; o_Bifidobacteriales; f_Bifidobacteriaceae; g_Bifidobacterium; s_bifidum
94290	80	83	86	91	83	p_Firmicutes; c_Bacilli; o_Lactobacillales; f_Streptococcaceae; g_Streptococcus; s_
99920	80	83	80	83	91	p_Firmicutes; c_Bacilli; o_Bacillales; f_Staphylococcaceae; g_Staphylococcus; s_
101886	37	37	43	43	43	p_Firmicutes; c_Bacilli; o_Lactobacillales; f_Streptococcaceae; g_Streptococcus; s_

M, mother's milk ecosystem; Opre, infant's oral ecosystem before breastfeeding; Opost, infant's oral ecosystem after breastfeeding; F, infant's fecal ecosystem.

Table A1. OTU sharing among the bacterial ecosystem of the infant's mouth and feces, and the mother's milk. The OTUs most frequently (>40% of the mother-infant pairs) shared by at least two bacterial ecosystems are shown, with assigned taxonomy. Only OTUs present at a relative abundance >0.1% were considered. Assigned taxonomy is reported using Greengenes syntax.

A PCoA based on unweighted UniFrac distance showed that the microbiota of infants' oral swabs, infants' feces and mothers' milk clustered separately (Figure A1.A), as expected being the resident communities of three distinct body districts that are different for pH, oxygen levels, and nutrients availability. Adonis test confirmed that the reported separation was significant, even if this result needs to be taken into account cautiously since the test for homogeneity of dispersion (function betadisper) returned that milk samples had a significantly different dispersion compared to the other groups of samples. When weighted UniFrac distances were used for PCoA (Figure A1.B), fecal samples overlapped with milk samples on PCo1. Thus, the difference between milk and fecal ecosystems was better explained by unweighted metrics, hinting that it might reside in fractions of the microbial communities that are exclusive of one of the two ecosystems (Lozupone *et al.*, 2007). Fecal samples also showed higher dispersion, indicating higher variability in the most abundant species of the ecosystem, with respect to oral and milk communities.

The oral microbiome was the least diverse of the considered ecosystems (Shannon diversity index, mean ± standard deviation (SD), 2.3±0.6; Figure A1.C), largely dominated by Streptococcaceae (average relative abundance (rel. ab.), 69.8%) (Figure A2), with *Streptococcus* being the dominant genus in 94% of samples, confirming the known literature on this ecosystem (Zaura *et al.*, 2014; Lif Holgerson *et al.*, 2015; Hendricks- Muñoz *et al.*, 2015; Davè *et al.*, 16). Also confirming the large amount of knowledge on the topic (Muller *et al.*, 2015; Jost *et al.*, 2012), the fecal microbiota of breastfed infants at 20 days of life was dominated by Bifidobacteriaceae (average rel. ab., 38.2%), with *Bifidobacterium* being the dominant genus in 67% of samples. Fecal microbiota also included relevant average abundances of Enterobacteriaceae (15.4%), Streptococcaceae (13.9%), Bacteroidaceae (9.5%), Staphylococcaceae (5.4%), and Lactobacillaceae (4.8%) (Figure 2).

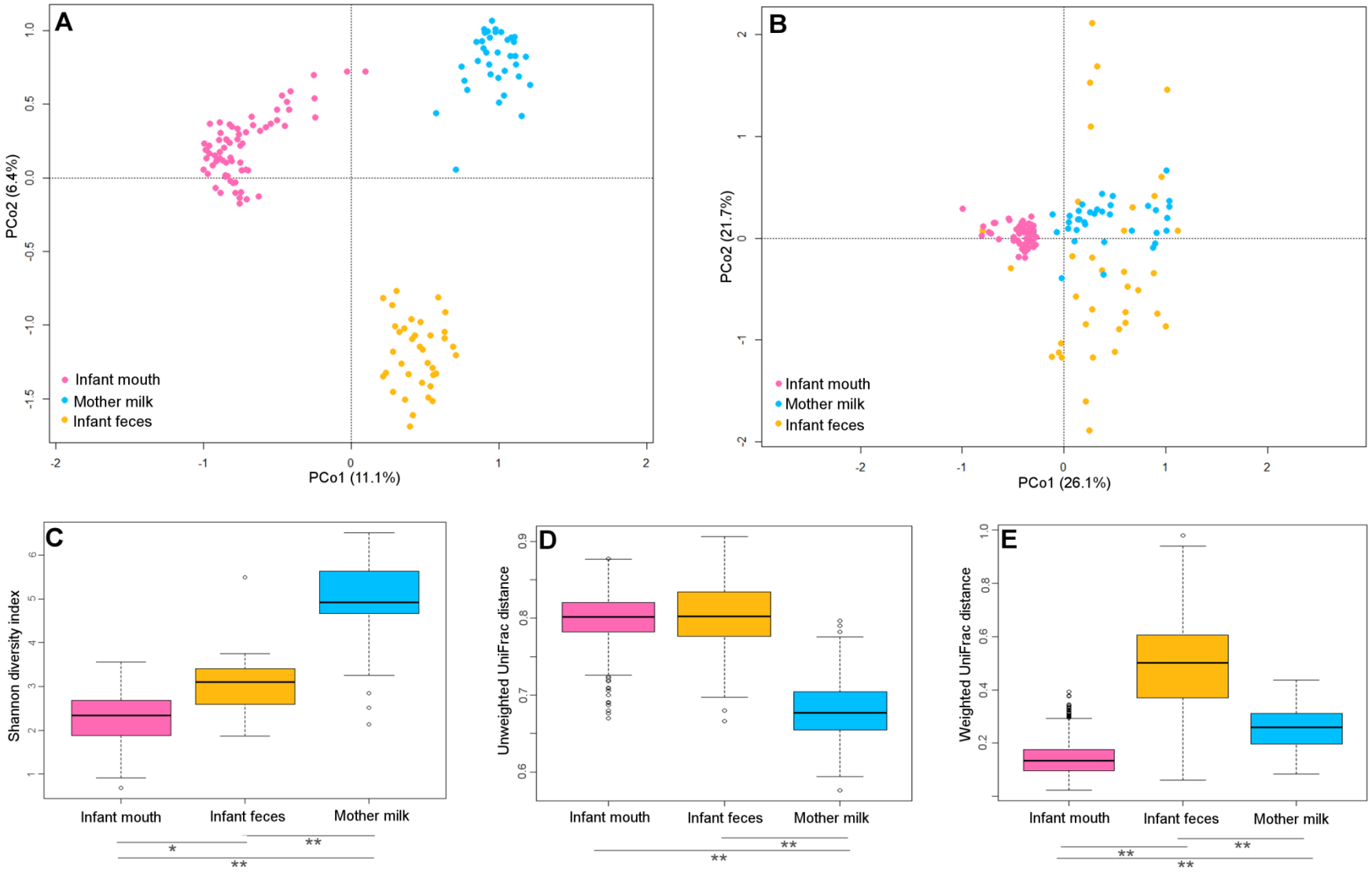

A
PCo2 (6.4%)
PCo1 (11.1%)
Infant mouth
Mother milk
Infant feces
B
PCo2 (21.7%)
PCo1 (26.1%)
Infant mouth
Mother milk
Infant feces
C
Shannon diversity index
Infant mouth
Infant feces
Mother milk
D
Unweighted UniFrac distance
Infant mouth
Infant feces
Mother milk
E
Weighted UniFrac distance
Infant mouth
Infant feces
Mother milk

Figure A1. Diversity in the bacterial ecosystem of the mother's milk, and infant's feces and mouth.
PCoA based on unweighted (A) and weighted (B) UniFrac distances of the microbiota of mother's milk (light blue), infant feces (yellow), and infants mouth (pink). Samples are identified by filled circles. In both PCoA first and second principal components (PCo1 and PCo2) are plotted. The percentage of variance in the dataset explained by each axis is reported. Box and whiskers distribution of the Shannon α-diversity index (C), intra-group unweighted UniFrac distances (D), and intra-group weighted UniFrac distances (E), calculated for milk (light blue), fecal (yellow) and oral (pink) samples. Significant differences between datasets are indicated, as calculated using Tukey post-hoc test after Kruskal-Wallis test for multiple comparisons (*, P≤0.001; **, P≤0.0001).

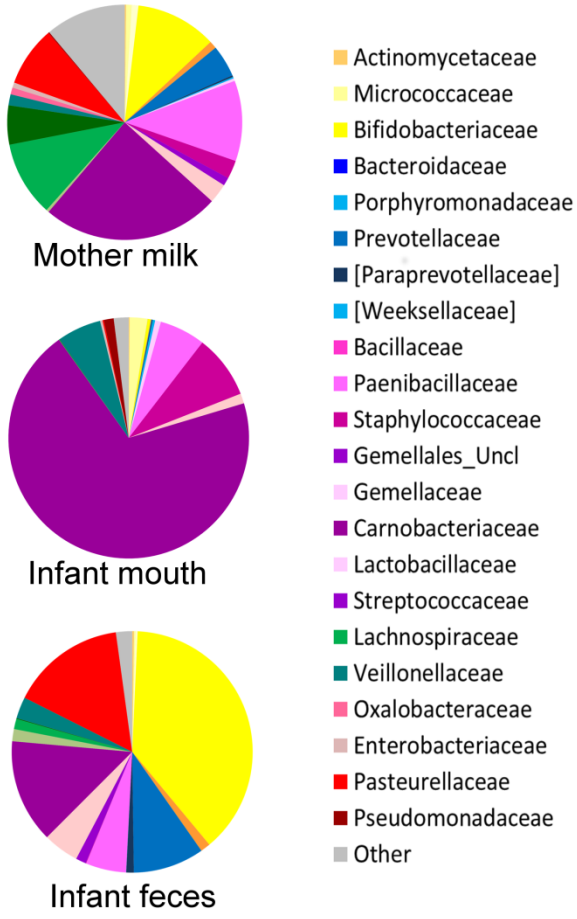


Figure A2. Average composition of the bacterial community in mother's milk and infant's feces and mouth. For each group of samples, a pie chart based on the average relative abundance (%) at family level is plotted. Bacterial families with relative abundance ≥0.2% in at least 10% of the samples are depicted. Colors for each family are reported in the legend.

Both oral and fecal microbiota of infants showed higher unweighted UniFrac distances within group (0.80±0.03 and 0.80±0.04, respectively) compared to milk (0.68±0.04, Kruskal-Wallis test, P<0.0001). When the weighted UniFrac metric was considered, the within-group distances obtained

for the fecal samples remained the highest (0.5±0.2), whereas those calculated for the oral samples became the lowest (0.14±0.06 Kruskal-Wallis test, P<0.0001). Since weighted UniFrac measure is better suited to detect differences in abundance even when the overall groups of organisms that are present in each sample remain the same (Lozupone *et al.*, 2007), these observations suggest that the variability between oral samples might reside in subdominant species that are not highly conserved among samples.

The average microbiota profile obtained for breast milk was significantly more diverse (Shannon diversity index = 4.9±1.1; Figure 1C) than both infant's feces and oral swabs (3.0±0.7 and 2.3±0.6, respectively; Kruskal-Wallis test, P<0.0001); interestingly, according to the unweighted UniFrac metric, the variability among milk samples was the lowest (0.68±0.03; Kruskal-Wallis test P<0.0001; Figure A1.D), and remained significantly lower than that of fecal samples when the weighted UniFrac distances were computed (0.25±0.07; P<0.0001; Figure A1.E). In other words, the milk ecosystem of the 36 enrolled mothers was richer and more similarly composed among samples (in terms of bacterial species) than the fecal or mouth ecosystem of their children, suggesting that the milk duct might act as an environmental filter allowing for the survival and proliferation of the same bacterial species in most individuals (a "niche-based" community assembly, according to Costello *et al.* (2012). As expected (Fitzstevens *et al.*, 2016), the milk ecosystem phylogenetic structure showed a slight dominance of Streptococcaceae (average rel. ab., 24.5%), with *Streptococcus* being the dominant genus in 53% of samples, but also a considerable representation of the typically infant fecal family Bifidobacteriaceae (11.2%, with *Bifidobacterium* being the dominant genus in 19% of samples) and Staphylococcaceae, which is instead a common skin and mouth inhabitant (Belkaid and Segre, 2014) (11.1%, with *Staphylococcus* being the dominant genus in 11% of samples) (Figure 2). Confirming previous studies (Jost *et al.*, 2014; Jost *et al.*, 2013) , and supporting the hypothesis of a possible link between the milk microbiota and the gut ecosystem of the mother, also anaerobic bacterial families that are commonly found in the adult human intestine, such as Lachnospiraceae, Ruminococcaceae and Bacteroidaceae, were present with an average rel. ab. of 10.3, 5.4 and 4.4%, respectively. Specifically designed studies are required to investigate if these bacteria are indeed alive in the milk ecosystem, as well as the ecological importance of this putative bacterial migration from the mother's gut to the milk duct. The milk microbiota was the only one showing a few genera that were present in more than half of the subjects but never retrieved in the other two ecosystems, partly confirming the observations of Jost *et al.* (2014). In particular, sequences assigned to the genus *Ralstonia* were detected at a relative abundance >0.1% in 83% of the milk samples (average rel. ab., 0.67%), and in none of the infant's fecal or oral samples; similar trends were shown by the genus *Sediminibacterium*

(detected in 61% of the milk samples, average rel. ab., 0.15%) and unclassified members of the Flavobacteriaceae family (detected in 50% of the milk samples, average rel. ab., 0.16%).

No correlation between the relative abundance of each family or genus in the three different ecosystems was found, after P values correction, with the exception of the abundance of the subdominant family Lactobacillaceae, whose values were positively correlated in saliva and feces of the same infant (Kendall tau = 0.61, P = 0.05).

According to our observations, the passage of the milk through the mouth affects the composition of the oral microbiota in each infant. Indeed, samples from the same subjects were rarely plotted closer to each other than to samples taken from other babies on the PCoA based on unweighted UniFrac distance (Figure A3), even if the multivariate analysis showed no significant separation between the two groups of samples (before and after breastfeeding). Significant differences were not found comparing the genus-level profiles of the samples before and after breastfeeding. However, it was possible to notice that 77% of the enrolled babies showed higher coordinate values on the PCo2 axis in the post-breastfeeding sample than in the pre-breastfeeding one. PCo2 was found to account for more variation in data than expected by random chance (broken-stick eigenvalue = 1.07, actual eigenvalue = 1.17). Indeed, the difference between pre- and post-breastfeeding PCo2 coordinates was found significant (paired Wilcoxon test, P = 0.02), meaning that it could be possible to find a common trend in the small changes occurring in the infant's mouth ecosystem after the mother's milk passage, in the frame of the individual microbiota structure. Pursuing the identification of these small changes, we found that the coordinate values on the PCo2 axis of each samples were significantly ($P<0.01$) and positively correlated to the relative abundance of a few bacterial families, which were found averagely more represented in the breast milk than in the infant's oral ecosystem, such as Lachnospiraceae (Kendall tau, 0.49; average rel. ab.: 0.07% in infant's mouth, 1.3% in infant's feces, and 10.3% in mother's milk), Ruminococcaceae (Kendall tau, 0.46; average rel. ab.: 0.04% in infant's mouth, 0.1% in infant's feces, and 5.4% in mother's milk), Oxalobacteriaceae (Kendall tau, 0.40; average rel. ab.: 0.29% in infant's mouth, 0% in infant's feces, and 0.72% in mother's milk), and Bacteroidaceae (Kendall tau, 0.48; average rel. ab.: 0.09% in infant's mouth, 9.5% in infant's feces, and 4.4% in mother's milk).

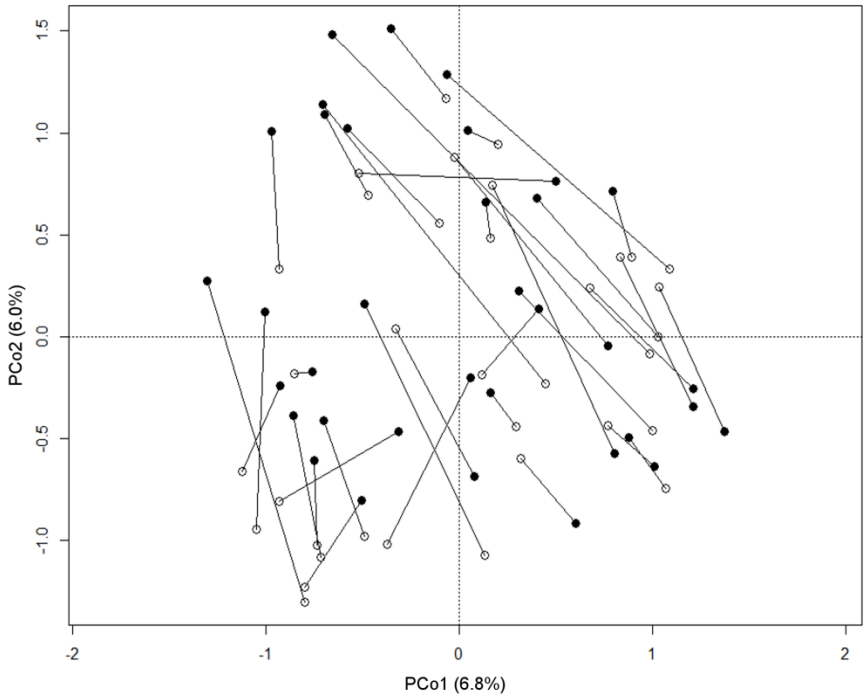


Figure A3. Relationship between pre- and post-breastfeeding infant oral microbiota. PCoA based on unweighted UniFrac distances of the microbiota of the infant's mouth sampled before (empty circles) and after (filled circles) breastfeeding. Samples from the same subject are connected by a black line. The first and second principal components (PCo1 and PCo2) are plotted. The percentage of variance in the dataset explained by each axis is reported.

Aiming at exploring the possibility of passage of bacteria from one ecosystem to another, we focused our attention to the OTUs shared between two or three samples taken from the same mother-infant pair. It is important to remember that 16S rRNA gene-based characterization does not allow for strain-level analysis. However, the sharing of the same OTUs might give indications on what species could be interesting to further explore (and possibly identify to the strain level) using culture-based techniques and/or metagenomics. Filtering for the OTUs accounting for at least 0.1% of the ecosystem diversity (number of normalized sequences per sample), a mean of 4.5 (considering pre-breastfeeding oral samples, range 2-10) and 4.7 (considering post-breastfeeding oral samples, range 1-11) OTUs were shared between the three ecosystems (Figure A4). Among those more frequently shared (Table 1) we found OTUs assigned to *Staphylococcus* spp. (shared by the three samples in 80% [pre-breastfeeding] and 83% [post-breastfeeding] of pairs, and by feces and milk only in 91% of cases), *Streptococcus* spp. (shared by the three samples in 80% [pre-breastfeeding] and 83% [post-breastfeeding] of pairs, and by feces and oral samples only in 86% [pre-breastfeeding] and 91% [post-breastfeeding] of cases), and *Streptococcus infantis* (shared by the three samples in 63% of pairs, and by feces and milk only in 97% of cases). Interestingly, the *Streptococcus* OTUs found to be preserved

among two or three ecosystems were also the dominant ones in all the infant's oral samples: indeed, one or a couple of these OTUs generally accounted for the totality of the Streptococcaceae population and, in most cases, for the dominant portion of the entire ecosystem, confirming previous findings on the adult's oral microbiota (Li *et al.*, 2013). Even if the genera *Streptococcus* and *Staphylococcus* have been recognized as universally predominant in the human milk by a recent systematic review (Fitzstevens *et al.*, 2016), the mechanisms of their colonization of the milk ducts are not explained. The very high abundance of *Streptococcus* in the baby's mouth that we report in the present study, together with the identity between the dominant *Streptococcus* OTUs in the infant's mouth and those detected in their mothers' milk, bring us to suggest that the infant's mouth could have a seeding effect on the milk duct resident community during suction.

Milk and infant's oral microbiota also shared the presence of a OTU assigned to unclassified members of Gemellaceae family, in 51% of mother-infant pairs; indeed, *Gemella* is another known major core genus in both adult and infant's oral mucosa (Costello *et al.*, 2013; Zaura *et al.*, 2014; Hendricks-Muñoz *et al.*, 2015).

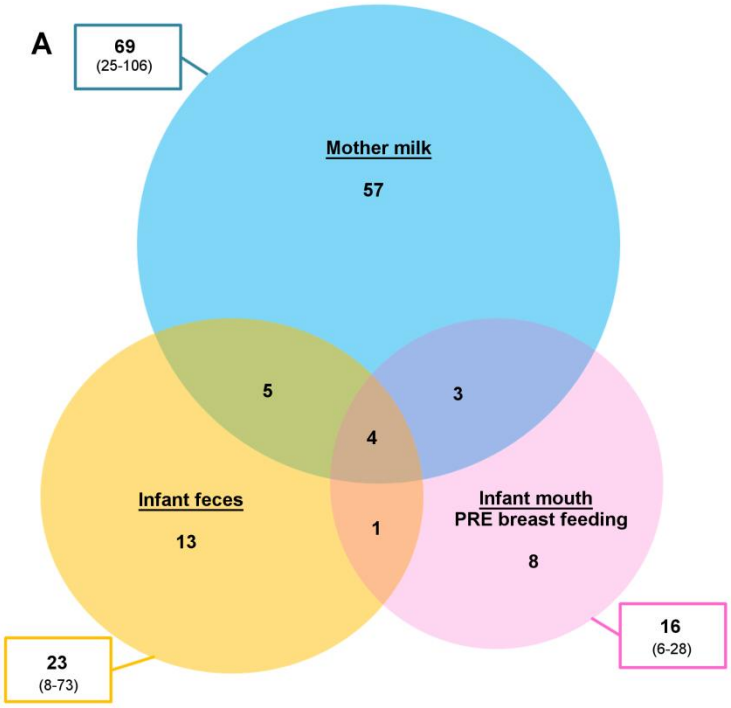
Most interestingly, the majority of the OTUs shared between mother's milk and infant's feces, but not present in infant's mouth, was assigned to members of the *Bifidobacterium* genus, well-known inhabitants of the gut microbiota of breastfed infants (Arrieta *et al.*, 2014) In particular, OTUs assigned to *Bifidobacterium breve*, *Bifidobacterium bifidum*, and *Bifidobacterium longum* were shared by 46%, 51%, and 74% of the milk and fecal samples taken from the same mother-infant pair (Table 1), supporting the hypothesis that the mother's milk may act as a reservoir of pioneer probiotic bacteria for the baby's gut microbiome (Jost *et al.*, 2013). These bacteria are necessary for the degradation of human milk oligosaccharides (HMO) and are boosted in the infant's gut by the continuous refueling of these energy source (Mueller *et al.*, 2015). It was not surprising to find that bifidobacteria were almost absent in the infant's oral ecosystem (average rel. ab., 0.4%), probably due to the aerobic environment provided by the baby's mouth; however, thanks to their known ability to tolerate oxygen exposure (Bottaccini *et al.*, 2014), they could be able to survive the transition through the oral cavity without actively colonizing it.

Our study has the limitations of a 16S-based molecular characterization, namely the possible biases deriving from the DNA extraction method, PCR amplification, and OTU assignment algorithm (Schloss and Westcott, 2011; Walker *et al.*, 2015), as well as the failure in discriminating between DNA deriving from live and actively proliferating bacteria and DNA fragments from dead cells. However, the homogeneity of our cohort, as well as the inclusion of oral samples from the infants before and after breastfeeding, led to interesting and useful observations that add knowledge to the complex, and still to be disentangled, topic of the microbiome assembly.

In particular, we observed a very limited number of shared OTUs and reported no correlation between the abundances of bacterial families or genera among the mother's milk, the child's mouth and the child's gut. These findings seem to support the hypothesis that, for most of the inhabiting species, the process of microbiota assembly in different infant's body sites and in the mothers' milk ducts is driven more by local adaptation than by true immigration of bacteria from other ecosystems, according to the metacommunity theory depicted by Costello *et al.* (2012).

On the contrary, an interesting behavior was observed for OTUs assigned to the genera *Bifidobacterium*, *Streptococcus* and *Staphylococcus*, which constitute a relevant fraction of the infant gut and mouth ecosystem. Indeed, among the OTUs assigned to these genera, a few were retrieved as dominant or very abundant in the majority of the infants and were also shared by the corresponding mother's milk. Even if the sharing of the same OTUs cannot be considered as a proof of transmission, the colonization of both the mother's milk and infant's feces by the same *Bifidobacterium* OTUs seems to sustain the hypothesis that the human milk is among the sources for the baby's gut inoculation of this bacterial group. At the same time, it does not constitute a proof that live bacteria can be translocated through an entero-mammary pathway (Rodríguez, 2014). Indeed, more recent observations (Meadow *et al.*, 2015) seem to imply that bacteria do not need to be transported through complex enteric mechanisms to migrate from one human ecosystem to other body sites or, possibly, to other individuals, but just to be "emitted" in the microbial cloud that surrounds each individual.

The very frequent retrieval of the same *Streptococcus* and *Staphylococcus* OTUs in the majority of the infants, as well as in their mothers' milk microbiota, is also an intriguing observation, because this consistency of behavior among the enrolled subjects might call for the existence of a biological or ecological role for these bacteria during the infant's microbiota assembly. A streptococcal and/or staphylococcal migration from one ecosystem to another cannot be proven by our results, but the very high abundance of *Streptococcus* spp. in the oral ecosystem leads us to speculate that the baby's mouth might be the among the sources of contamination of both the infant's gut ecosystem, via deglutition, and mother's milk ducts, during suction. This will need to be proven by cultivation-based studies, where strains can be isolated and fully characterized, as well as by studies with a longitudinal layout for all the considered ecosystems. Moreover, since all the enrolled infant were born in the same hospital, it cannot be excluded that the frequent and abundant retrieval of the same *Streptococcus* and *Staphylococcus* OTUs might be linked to the contact with the same environment in the very first days of life; this observation could yet strengthen the importance of the living environment in determining the human microbiome composition.



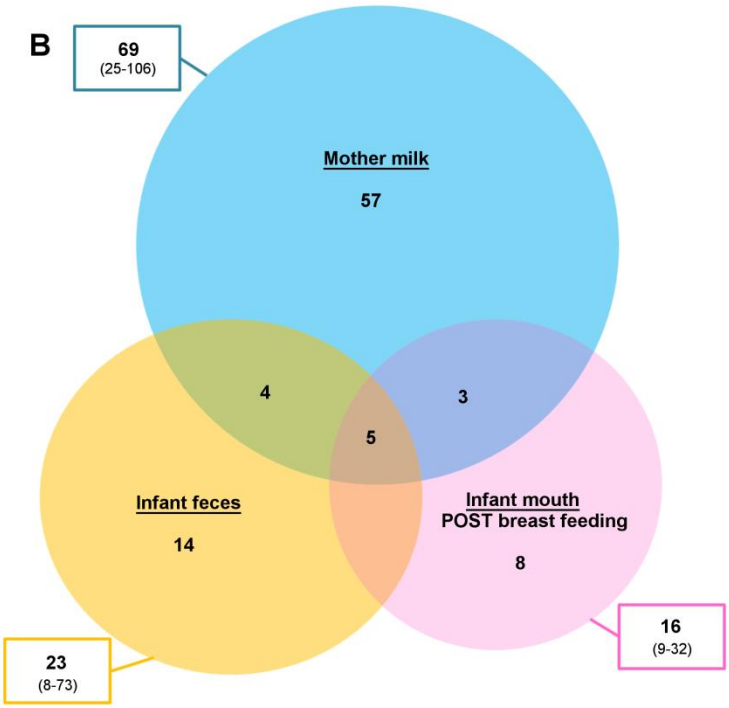


Figure A4. OTU sharing between mother's milk, and infant's fecal and oral bacterial ecosystems. Venn diagrams showing the average number of OTUs shared between the bacterial communities of the mother's milk (light blue), the

infant's feces (yellow) and the infant's mouth (pink), the latter sampled before **(A)** and after **(B)** breastfeeding. The total number of OTUs for each ecosystem is reported in the boxes outside the circles, expressed as mean and range (in brackets).

Conclusions

The assembly dynamics of the infant's gut ecosystem are a topic of huge interest for human immunology and microbiology (Lynch and Pedersen, 2016). Indeed, the existence of a crucial window of time in which the microbiota contributes to the education of the infant's immune system has been demonstrated (Arrieta *et al.*, 2015; Honda and Littman, 2016). Our study highlights that bacterial communities in other body sites could be involved in the early phases of the gut microbiota assembly. Even if the specific conditions (pH, oxygen level, nutrient availability) of the infant's gut seem to be the most relevant filter impacting on its final phylogenetic structure and abundance profile, other bacteria-colonized districts and/or the bacteria-coated body surfaces of the mother might act as a reservoir of seeding species, among which those ecologically necessary and intestinally-adaptable will be selected.

Part B: Dynamics of the infant microbiomes onset: exploring the gut and the oral microbiota in full-term and pre-term infants in the frame of mother's milk microbial ecosystem

Introduction

The early gut microbial community establishment in infants during delivery and the very first days after birth has been shown to contribute in building a solid healthy status for the following age later in life (Arrieta *et al.*, 2014). The gestational age at birth is one of the first factor that affects the intestinal colonization (Gregory *et al.*, 2016, Microbiome). Moreover, sooner the infant is born, higher is the risk of complication onset. Normal duration of pregnancy settles on 40 weeks, but infants born between 37 and 40 weeks are in any case considered full-term. For what concerns preterm birth, the World Health Organization has drawn up a classification based on gestational age:

- extremely preterm (<28 weeks);
- very preterm (28 to <32 weeks);
- moderate to late preterm (32 to <37 weeks);

In 2012, in the face of 131,296,785 global births, 15,000,000 infants were born preterm. Of these, 84.3% belong to the category "moderate to late preterm" (OMS report, Born to soon, 2012). Preterm births can be provoked by maternal or foetal infections and health conditions of the mother (ascribable to nutrition, psychological stress or anxiety, smoking, age) (Ruiz *et al.*, 2016). The common procedure for delivery in preterm birth is C-section, in order to avoid the stress linked to the delivery process to the foetus (Italian Society of Gynecologics and Ostetricians, Gestione del parto pretermine). C-section leaves a very peculiar signature on gut microbial community, characterized by a higher relative abundance of *Clostridium*, *Staphylococcus*, *Propionibacterium* and *Corynebacterium* with respect to vaginally-delivered infants (Dominguez-Bello *et al.*, 2010). While the gut microbial colonization in full-term infants has been extensively investigated (Jost *et al.*, 2012; Arrieta *et al.*, 2015; Honda and Littman, 2016; Lynch and Pedersen, 2016), focus has been moved also on the preterm ones only recently (Gregory *et al.*, 2016; Arboleya *et al.*, 2011; Arboleya *et al.*, 2017). Among the 3 categories of preterm infants, the moderately to late preterm one is the less examined, although it is the most frequent type of premature birth in the world. These infants are subjected to many complications, (such as necrotising enterocolitis and sepsis) and, even their clinical history is often less complex compared to the extremely and very preterm ones, they have to be followed carefully, because they often develop more critical conditions than full-term babies. Interestingly, the GM has been shown to be implicated in the onset of these complications (Pammi *et*

al., 2017). The gut microbial community of moderately to late preterm babies has been analysed through 16S rDNA Illumina sequencing in a few papers (Gregory *et al.*, 2016; Sherman *et al.*, 2016), with a limited sample size (Berrington *et al.*, 2013). The additional obstacle to these studies is the tremendous number of variables affecting the clinical history of these subjects. Indeed, they are often administered with antibiotics and probiotics, together with surfactants. Moreover, until the infant cannot feed itself from the maternal breast, she/he is nourished using enteral nutrition, with a mixture of maternal breast milk (when available), human breast milk from donor and formula. Breast milk is considered the best choice for infants' health and it contributes to GM shaping not only as a source of prebiotic nutrients (above all human milk oligosaccharides, HMOs), but also as a putative supply of bacterial cells (especially microorganisms belonging to the families of Streptococcaceae, Paenibacillaceae, Lachnospiraceae, Bifidobacteriaceae and Pasteurellaceae) (Biagi *et al.*, 2017). These bacteria maybe are among the first colonizers of infant's gut, but the vertical transmission mechanism is still scarcely depicted, especially considering that also mother's skin microbiota and infant's oral microbial community are involved in this process. Infant oral cavity is a forced route for the milk to cross during feeding. Moreover, during suction, it also acts as collector of the mother's skin microbial cells. The oral microbiota has been extensively described and we know that healthy infant salivary community is constituted mainly by microorganisms belonging to the family of the Streptococcaceae (with *Streptococcus* being the dominant genus), Staphylococcaceae, Paenibacillaceae and Veillonellaceae (Biagi *et al.*, 2017).

In order to shed light in the intestinal colonization process in moderately to late preterm infants, we characterize through Illumina sequencing stool sample collected from 21 infants, besides oral swab samples and milk samples from the mothers, having, in our knowledge for the first time, a highly comprehensive vision of the bacterial ecosystems that surround this neglected preterm category.

Subjects recruitment

Sixteen mother-infant pairs (16 mothers and 21 infants) were recruited in the Nursery of S. Orsola-Malpighi Hospital in Bologna, Italy. Five of the 16 recruited infants were twin pairs, of which 1 pair monochorionic monoamniotic and 4 pairs monochorionic biamniotic. Inclusion criteria matched the WHO definition for moderate to late preterm, according to which gestational age must be included between 32 and 37 weeks. Exclusion criteria were the necessity or preference of the mother to use formula milk instead of human breast one (both maternal and from donor). Anthropometric and clinical characteristics of the infants recruited are reported in Table B1. Milk samples from the

mother, faecal samples and oral swabs from the infant were collected at birth and at 4, 7, 14, 21 and 30 days after birth.

Faeces were collected from diapers using a standard sterile collection tube. Milk samples were collected with the aid of a breast pump into sterile plastic tubes; prior to collection, mothers were asked to wash the nipple and mammary areola with soap and water. Oral samples were obtained by gently swabbing a sterile cotton-tipped applicator on the inside of the infant's cheek. All samples were immediately delivered to the laboratory and stored at -80°C until analyses. Demographic and clinical data were recorded in a specific case report form. All participants signed a written consent form. The study was approved by the ethics committee of the S. Orsola Malpighi Hospital in Bologna (study protocol 53/2014/U/Tess). The study was conducted according to the principles expressed in the Declaration of Helsinki. Methods were carried out in accordance with the approved guidelines.

Gestational age at delivery	33.4±1.0
Weight at birth (g)	1787.3±444.9
Female and male	13/21 (61.9%) and 8/21 (38.1%)
Vaginal delivery	2/21 (9.5%)
C-section	19/21 (90.4%)
Twin	10/21 (47.6%)
Sepsis	4/21 (19.0%)
NEC	2/21 (9.5%)
Probiotic administration (*Reuflor*)	17/21 (80.9%)
Surfactant administration	3/21 (14.2%)
Antibiotic administration	14/21 (66.6%)
of which ampicillin	11/14 (78.6%)
of which ampicillin+amikacin	3/14 (21.4%)
Antimycotic administration (fluconazole)	1/21 (4.8%)
Enteral feeding	21/21 (100%)
Range start enteral feeding (days)	1-5 (mean 1.4± SD 1)
Feeding	
maternal breast milk+ formula	12/21 (57.1%)
human breast milk from donor+formula	1/21 (4.7%)
human breast milk from donor+maternal breast milk	6/21 (28.5%)
human breast milk from donor	2/21 (9.5%)
Breastfeeding	15/21 (71.4%)
Range start breastfeeding (days)	7-45 (mean 16.6 ± SD 9.9)
Demission (temporal range in days)	7-45 (mean 20.1 ± SD 10.4)

Table B1. Anthropometric and clinical characteristics of the pairs mother-infant recruited. Descriptive characteristics of pre-term infants, together with the drug therapy, probiotics administration and the feeding type are reported for each subject recruited (21 infants and 16 mothers)

Experimental procedure and statistics

Samples were processed as described in "Experimental procedure". Statistics was performed using RStudio software version 1.0.136 running on R software 3.1.3 (https://www.r-project.org/), implemented with the libraries vegan and made4. Relative abundance filtering was performed keeping the genera showing a minimum abundance of 2% in at least 10% of the samples. Weighted and unweighted UniFrac and Bray-Curtis distances were used for Principal Coordinates Analyses (PCoA), and the significance of separation was tested by permutational multivariate analysis of variance using the function "Adonis" of the vegan package. Wilcoxon test was used to assess significant differences between two groups of samples. Correlation between datasets was verified computing Kendall correlation coefficient.

Results and discussion

In this research work we characterized the GM community and the salivary microbiota in 21 moderate to late preterm infants, from birth until the 30th day of life, through the Illumina sequencing of 16S rDNA. Moreover, the microbial community of the maternal milk (when available) was described. We sequenced a total of 348 sample (138 stools, 69 milk samples and 141 salivary samples), obtaining a number of 11,810,468 high quality reads with a mean of 66,713.2069 ± SD 271,834.7874 reads per sample. Reads were clustered in 46,404 operational taxonomic units (OTUs) at 97% of identity.

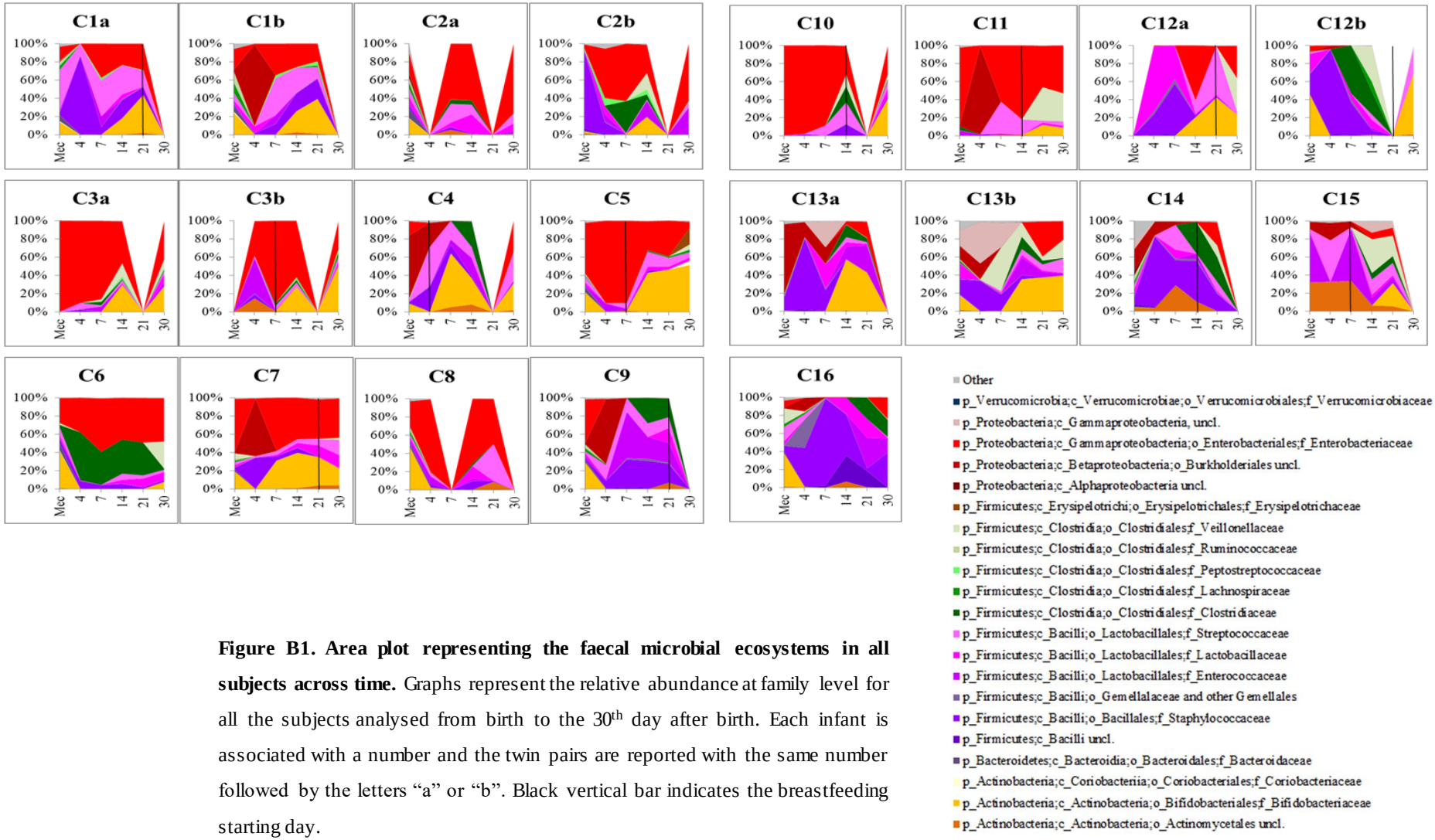


Figure B1. Area plot representing the faecal microbial ecosystems in all subjects across time. Graphs represent the relative abundance at family level for all the subjects analysed from birth to the 30th day after birth. Each infant is associated with a number and the twin pairs are reported with the same number followed by the letters “a” or “b”. Black vertical bar indicates the breastfeeding starting day.

When the community structure of all the stool samples collected was analysed at family level, we observed that faecal microbiota composition was extremely variable over time and between subjects, probably reflecting the personal clinical history of these infants (Figure B1).

The gut microbial community of the infants was analysed considering all the samples collected in the first 30 days after birth, excluding the meconia. Meconia are indeed a foetal product, composed by bile, mucus, epithelial cells and swallowed amniotic fluid. During the first month of life the infant's GM was constituted mainly by three families: Enterobacteriaceae, Staphylococcaceae and Bifidobacteriaceae. When this fraction is described at genus level, we observed that there is a core community of bacteria that accounts for 42.7% of the ecosystem, and this core is averagely composed by *Klebsiella* (17.1%), *Staphylococcus* (13.2%) and *Bifidobacterium* (12.3%). These results partially resemble the ones reported by Stewart and colleagues (2016) on pre-term GM composition. One infant (C6) has a very peculiar GM structure, that doesn't resemble any of the other infant's microbial community. Indeed, it is characterized by a high level of Clostridiaceae starting from day 4 after birth. When the gut microbial composition was observed at deeper phylogenetic level (namely species), we found out that the total amount of Clostridiaceae is due to *Clostridium neonatale*. This species accounts totally for the Clostridium fraction until day 14, then at day 21 *Clostridium perfringens* appears. At day 30 their relative abundance decreases dramatically, giving space for news *Clostridium* species to proliferate (*Veillonella dispar* and *parvula*). In this infant, a pattern very similar to the Clostridiaceae one is followed by the family of Enterobacteriaceae. Indeed, only one genus accounts for the total amount of this family, which is *Citrobacter*. At day 21, another member of the Enterobacteriaceae family appears, that is *Klebsiella*. Both members of the Enterobacteriaceae family have been retrieved in stool samples of infants who are hospitalized in intensive care unit (Goldman *et al.*, 1978). The distinctive tract of this baby in respect to the other is the administration of fluconazole, that was given from day 1 to day 10. The impact of fluconazole on infant GM has never been investigated yet, but a study has been conducted in healthy adults (Wheeler *et al.*, 2016). The response to this antimycotic treatment of the 2 microbial communities appears to be diverse, but we must take into account the profound differences between infant's and adult's GM.

When we described the gut microbial ecosystem trend in terms of biodiversity (expressed by the Shannon index) over time, we noticed that, at birth, the meconium was highly biodiverse (mean 3.68 ± SD 1.64) and then passed through a drop at day 4 of life (mean 1.86 ± SD 0.90). This variation was significant (Wilcoxon rank sum test p=0.0002). At day 7 the biodiversity started to grow (mean 2.83 ± SD 0.89) until reaching a plateau at day 14 (mean 3.48 ± SD 0.94). Interestingly, the same trajectory was followed by the salivary microbiota, thus maintaining always a lower biodiversity level compared to stool samples (Figure B2). Also the salivary microbiota alpha diversity followed a drop from day

1 to day 4 after birth and also in this case the variation was significant (Wilcoxon rank sum test p=0.0002).

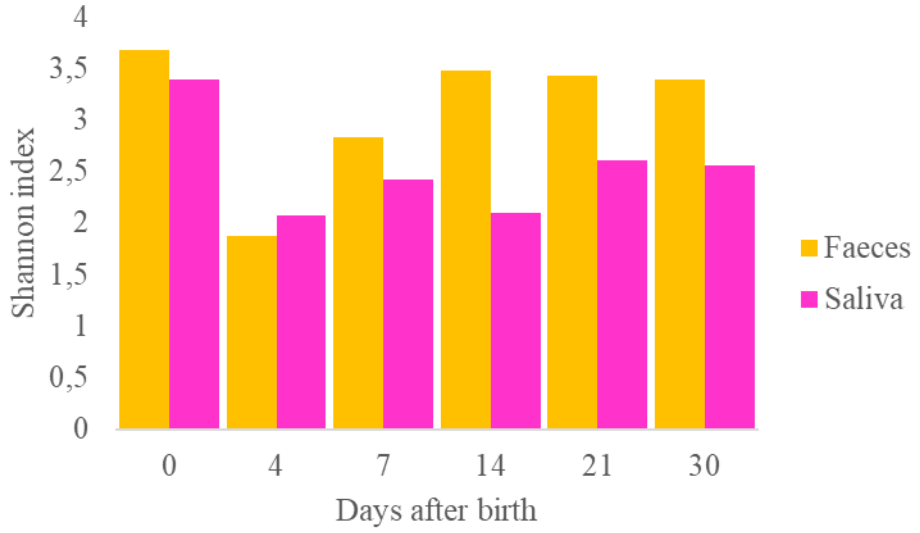


Figure B2. Bar plot representing the alpha diversity of stool and saliva samples over time. In the graph the mean alpha diversity measured by Shannon index of the stool and saliva samples divided according to the timepoint of collection is represented. Yellow bars represent stool samples and pink bars the salivary microbiota.

A PCoA based on Bray-Curtis distance showed that the faecal samples collected from day 4 to day 30 after birth cluster separately according to timepoint (Figure B3). This separation was confirmed by permutation test with pseudo F ratios ($p<0.001$). Stool samples also describe a trajectory along timepoints: we noticed that, as the infants grow older, faecal samples move from the down right of the graph to the above left side. Interestingly, when we correlated the bacterial genera with the PCoA axes, we noticed that MDS2 axis is positively correlated with *Bifidobacterium* ($p<0.001$), *Delftia* ($p<0.001$), *Faecalibacterium* ($p<0.001$), *Lactobacillus* ($p<0.001$) and *Roseburia* ($p=0.001$). These bacteria (*Delftia*, *Faecalibacterium* and *Roseburia*) are common commensals of the human adult gut ecosystem, pointing that the infant GM is approaching to an adult-type community as the baby grows older. The positive correlation with *Bifidobacterium* may indicate the increasing chances of the infant to be breastfed. In our cohort, the infants started to be breastfed between 7th and 45th day (mean 16.6 ± SD 9.9).

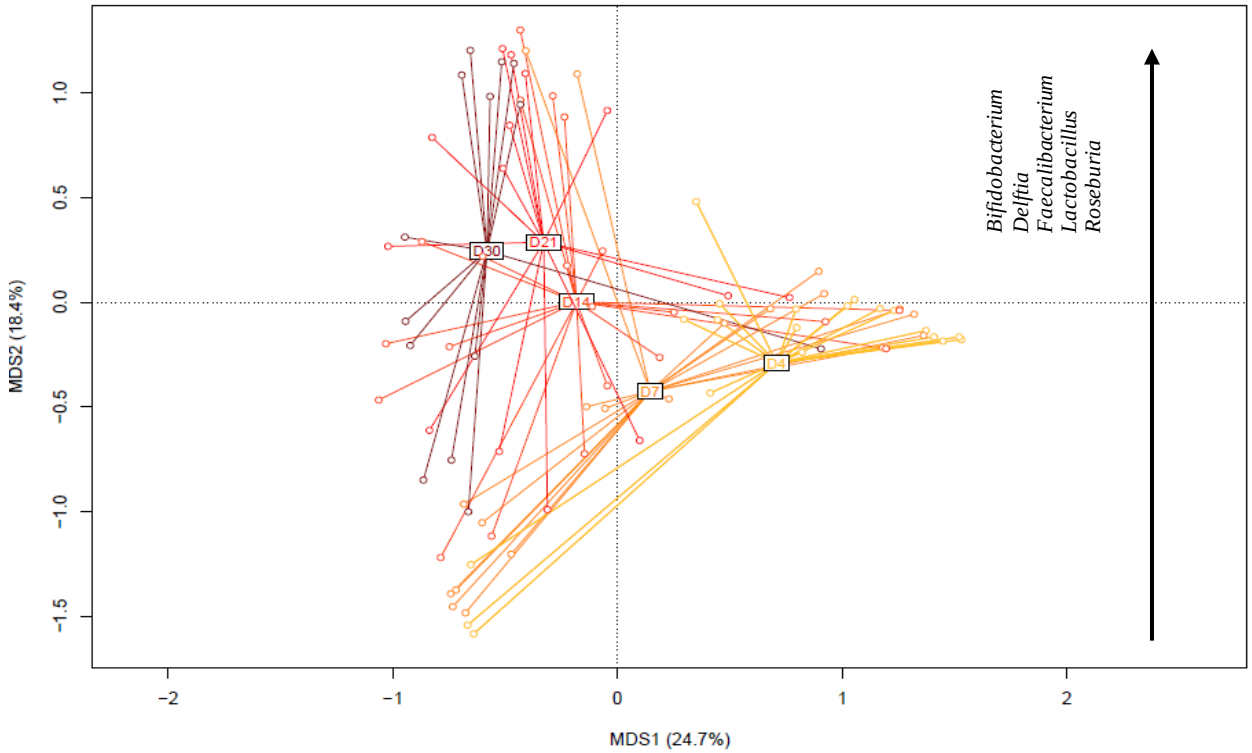


Figure B3. Principal Coordinates Analysis based on Bray-Curtis distances of faecal samples from day 4 to day 30 after birth. All faecal samples collected belonging to the same timepoint are represented with the same colour. Axis MDS1 and MDS2 account for the 24.7% and 18.4% of the variability respectively. When the separation between timepoints is measured, the result is significant (permutation test with pseudo F ratios, $p<0.001$). Based on the sample ordination, significant positive correlations between MDS2 and the relative abundance of some intestinal bacterial genera are reported.

The Bray-Curtis average distance at the same timepoint for stool samples increases from birth to day 14 (0.65 to 0.77), suggesting an impact of the clinical conditions on the gut microbial composition. Starting from day 21 (0.54), the distance decreases, indicating that stool samples composition is more similar. It's tempting to suggest that the GM structure follows an adaptive trajectory towards a healthy-like infant composition. The same path is followed by the saliva samples, for which interindividual diversity increases over time from birth to day 14 of life, to then start falling at day 21. Intriguingly, the Bray-Curtis average distance between samples belonging to the gut ecosystem, regardless to timepoint, is higher compared to the salivary samples one (Figure B4).

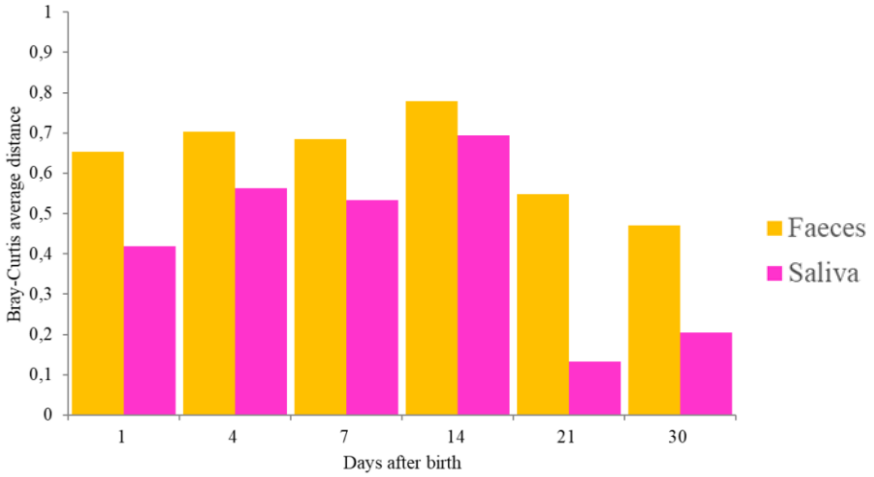


Figure B4. Bar plot representing the Bray-Curtis average distance of stool and saliva samples over time. In the graph the distance between faecal and saliva samples belonging to the same ecosystem collected at the same timepoint is measured by Bray-Curtis metrics. Yellow bars represent faecal microbiota and pink bars the saliva microbiota.

As reported before, meconium samples were analysed separately as they are a foetal product. When the bacterial diversity between meconia was measured using the unweighted UniFrac distances and represented as PCoA, we noticed that a separation emerged. While the samples C1a, C1b, C2a, C8, C14, C9, C13b, C5, C7, C6, C4, C13a, C2b, C16 and C12b clustered on the right side of the graph (group A), C11, C10, C3a and C15 were spread on the left side. The separation between group A and the outliers is significant (permutation test with pseudo F ratios, p=0.002) (Figure B5).

A

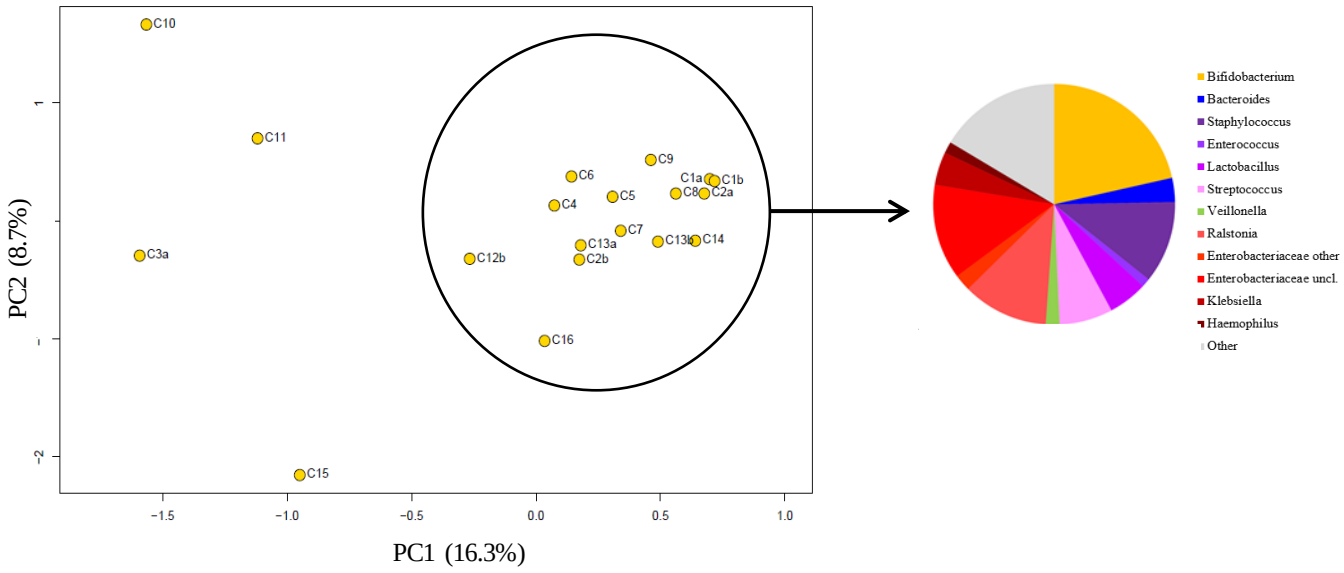


Figure B5. Principal Coordinates Analysis based on unweighted UniFrac distances of meconium samples. Meconium samples analysed are plotted based on the unweighted UniFrac distances. Axis PC1 and PC2 account for the 16.3% and 8.7% of the variability respectively. The formation of a group appears (group A) and when the difference between this group and the outliers is measured with permutation test with pseudo F ratios, the p values result is significant (p=0.002). The average GM composition of samples belonging to group A is represented with pie chart.

The same separation was confirmed when a heatmap was built on the relative abundance at genus level, using Spearman distance and Ward clustering method (Figure B6).

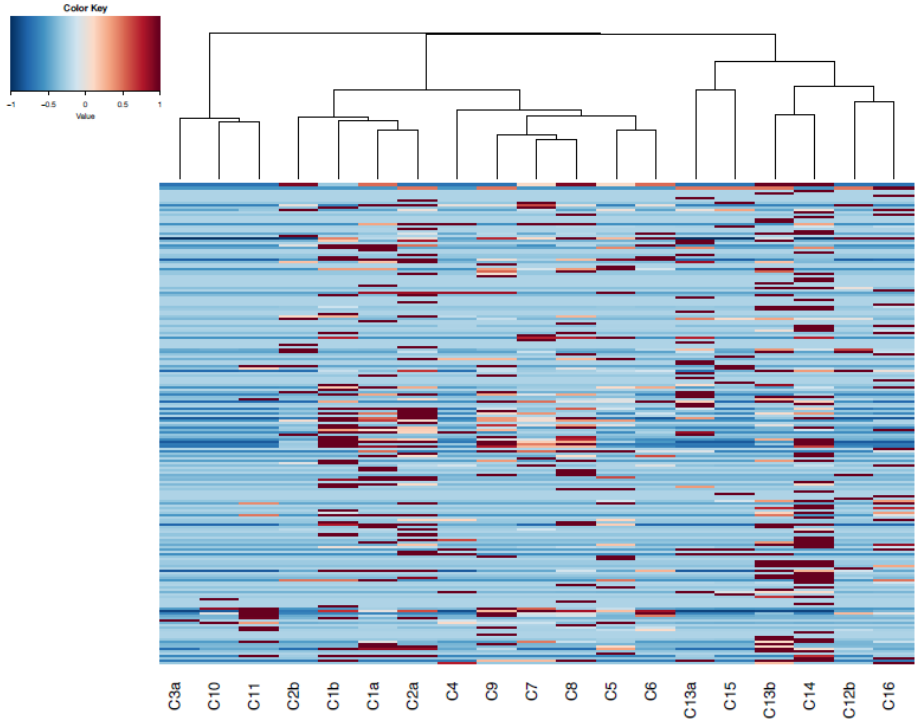


Figure B6. Heat map based on genera composition of meconium samples. Heat map shows the relative abundance of meconium composition. Hierarchical clustering was performed using the Pearson distance measure and Ward linkage method.

Taking into account these results, we proceeded to mediate the meconium relative abundance at genus level of those subject belonging to the group A and obtained a community characterized by 21.4% of *Bifidobacterium*, 11.4% *Ralstonia*, 11% *Staphylococcus*, 7.1% *Streptococcus*, 5.2% *Lactobacillus*, 4.4% *Klebsiella*, 3.2% *Bacteroides* and 12.7% of unclassified bacteria belonging to the family of Enterobacteriaceae (Figure B5). According to Chu and colleagues (2017), meconium structure at genus level (measured for the infants recruited with the same gestational age we are considering) is mainly characterized by 13.9% of *Sphingobium*, 9.9% *Neisseria*, 9.8% *Lactobacillus*, 8.9% *Staphylococcus* and 8.2% *Escherichia*. This meconium structure resembles only partially the one we

observed, and it's interesting to notice that no bifidobacteria was detected by them. On the other hand, outlier samples are characterized by a very individual composition, often presenting a dominance of few bacterial genera. C3a GM is composed by the 99% of *Erwinia*, C10 by 89% of *Klebsiella*, C11 by 57.61% of unclassified member of the Enterobacteriaceae, 16.77% of *Klebsiella*, 6.83% of *Citrobacter* and 3.45% of *Trabulsiella* and C15 by 54% of *Enterococcus* and 31% of *Rothia*.

What leaps out is that, while meconium structure appeared to be highly diverse, the alpha-diversity falls at the subsequent timepoint (day 4), while the inter-individual variation increased. This observation is supported by the comparison of the measure of the alpha diversity in meconium and stools collected at day 4, which results in a significant variation (Wilcoxon rank sum test, $p < 0.001$). Analogously to faecal samples, the milk samples presented an extremely variable composition over time and among subjects (Figure B7). Staphylococcaceae is the family that dominate the ecosystem, likewise reported in Biagi *et al.* (2017). Staphylococcaceae, in particular genus *Staphylococcus*, is a common symbiont of skin microbiota and it also retrieved in infant oral environment (Grice *et al.*, 2011; Biagi *et al.*, 2017). This suggests that mother's skin and infant's oral microbiota could act as seeding ecosystems during suction, taking advantage of the temporary opening of milk ducts. On the other hand, we can't exclude the possibility that, during sampling, skin microorganisms have been picked up erroneously.

We kept on examining the impact of breastfeeding on the biodiversity of the ecosystems focusing on the milk microbial community. We compared the alpha diversity values (measured with Shannon index) of the samples available before and after breastfeeding, but the result was not significant (Wilcoxon rank sum test, $p>0.05$), indicating that suction doesn't have any impact on the milk microbiota in terms of biodiversity. We performed the same analysis on stool samples, comparing alpha diversity values measured with Shannon index of the samples before and after the start of breastfeeding. The difference was significant (Wilcoxon rank sum test, $p<0.05$), probably due to the introduction of milk and the concomitant contact with the mother skin that happens during suction, that could act as an additional source of bacteria.

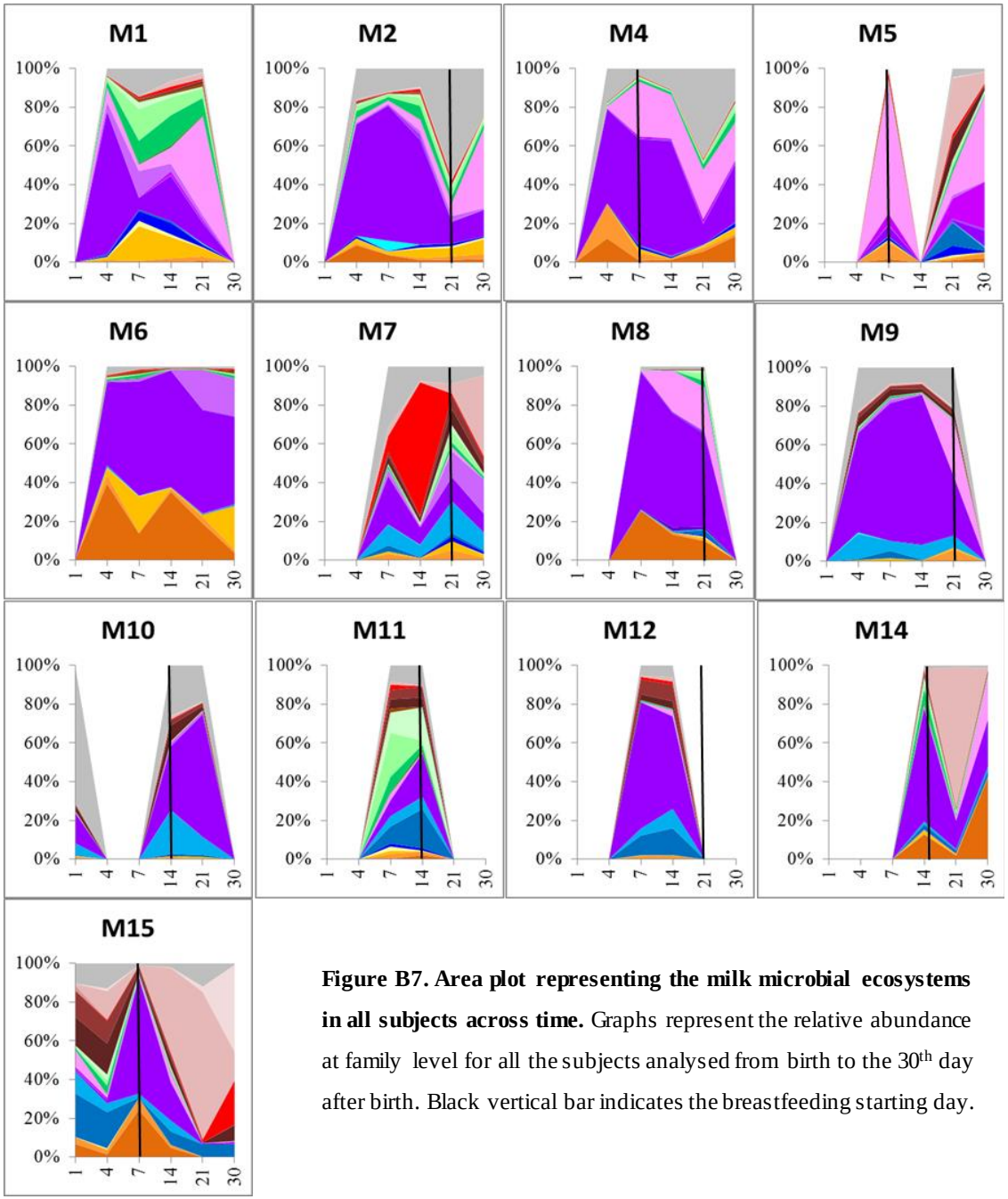


Figure B7. Area plot representing the milk microbial ecosystems in all subjects across time. Graphs represent the relative abundance at family level for all the subjects analysed from birth to the 30th day after birth. Black vertical bar indicates the breastfeeding starting day.

The salivary microbiota, differently from stool and milk samples, follows a reproducible pattern in terms of composition and dynamics between subjects. For this reason, we chose to mediate the relative abundances of the families composing the saliva microbial community of all subjects at the same timepoint and represent them as a dynamics over time, from birth to 30th day of life (Figure B8). We observed that Pseudomonadaceae, Oxalobacteriaceae and Straphylococcaceae decrease over time, while Streptococcaceae and Micrococcaceae (which are typical residents of the salivary microbiota) increase. This may indicate that, as the infant grows older, its oral microbiome changes and starts resembling the adult-type one.

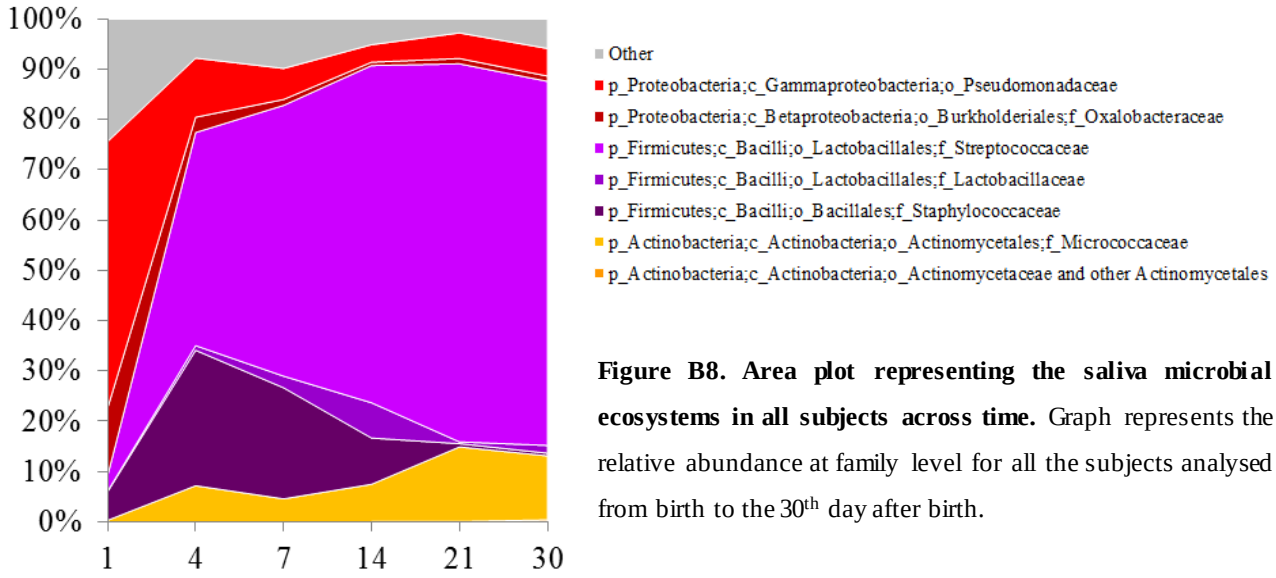


Figure B8. Area plot representing the saliva microbial ecosystems in all subjects across time. Graph represents the relative abundance at family level for all the subjects analysed from birth to the 30th day after birth.

We investigated if the start of breastfeeding had an impact also on oral microbiota composition, comparing the alpha diversity values of the samples before and after breastfeeding, measured by Shannon index, but the result was not significant (Wilcoxon rank sum test, $p>0.05$). Indeed, we assessed that the beginning of breastfeeding influences the intra-individual biodiversity only in stool samples, while it has no significant impact on mother milk and infant oral microbiota. This could suggest that the microorganisms that the infant acquires during breastfeeding, cross only transiently the mouth, without influencing the already resident community. This could be due to the differences in terms of environment that characterize mouth and gut (pH, availability of nutrients, oxygen level). However, once reached the gut, the same microorganisms are able to impact the GM composition.

Comparison between full-term and moderately to late pre-term

The characterization of the GM in moderately to late pre-term infants it is very important in the frame of depicting the role of GM in the onset of complications typical of pre-term infants, such as NEC

and sepsis. Furthermore, it's crucial, in a view of developing a personalised strategy of intervention, to understand the similarities and the differences with a heathy-like GM composition. For this reason, we proceeded to perform the comparison between gut microbial community in full-term and moderately to late pre-term. We thus compared the faecal GM structure of the 2 cohorts at 21 and 20 days of life respectively. These new-borns are exposed to different conditions, not only in term of gestational age, type of delivery and nutrition. Moderate to late pre-term infants are indeed often hospitalized in intensive care unit, administered drugs and fed with feeding tube. In our cohort, 90.4% of the infants were born with C-section, they remained in hospital from 5 to 45 days, the 71.4% of them was administered antibiotic and all of them were initially fed though feeding tube (Table B1). Considering that a high biodiversity of the ecosystem is positively correlated with higher ability of the GM to respond to external stressors, we evaluated firstly the alpha diversity (measured with Chao1 index) of the samples, comparing full-term babies and pre-term babies at 20 and 21 days respectively for all the ecosystems analysed. While milk samples showed a similar value of diversity, with no significant difference between full-term and pre-term infants (Wilcoxon rank sum test, $p=0.8$), a significant difference was found for faecal and oral microbial communities (Wilcoxon rank sum test, $p<0.001$ for faeces and $p=0.005$ for saliva). (Figure B9). The pre-term infant's ecosystems resulted to be more diverse, but with a concomitant broader interval of dispersion of the data.

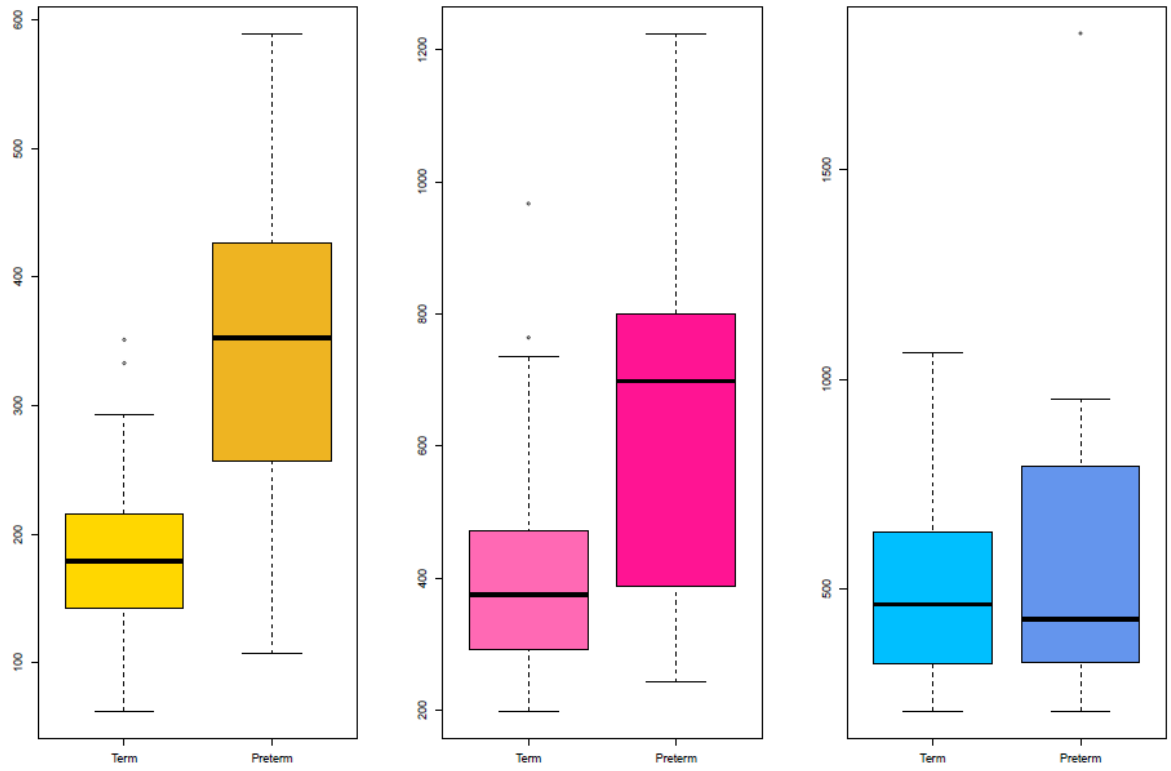


Figure B9. Boxplot representing the alpha-diversity of the 3 ecosystems (stool, saliva and milk) of term and preterm infants. Alpha diversity measured with Chao1 index is compared between full-term and moderately to late pre-term infants. The difference between the 2 cohorts is significant for faecal (yellow) and oral (pink) ecosystems (Wilcoxon rank sum test, $p<0.001$ for faeces and $p=0.005$ for saliva), while the difference is not significant for milk (azure) sample (Wilcoxon rank sum test, $p>0.8$).

These data suggest that gestational age of the infant, together with the clinical conditions, affect the alpha diversity of stool and saliva microbial community, but has no significant impact on the milk ecosystem.

In order to examine if samples belonging to the same ecosystem differ between the two cohorts analysed, a PCoA for stools, saliva and milk was performed using Bray-Curtis metrics (Figure B10). When the evaluation of the separation between the two cohorts was calculated using Adonis permutational test, the results were significant for all the three ecosystems (stool, $p=0.001$; milk, $p<0.001$; saliva, $p<0.001$). What leaped out was that, for faeces and milk samples, the dispersion was higher in pre-term infants than full-term ones. On the other hand, salivary microbiota samples in pre-term babies clustered very closely, indicating a very high similarity degree, maybe due to the restrict environment they are exposed to. Indeed, full-term babies at 20th day of life were not only breastfed, but also permanently residing at home and therefore are exposed to many contaminations, in terms

of environment and people. Instead, for what concerns our pre-term cohort, only 33% of the infants at 21st day of life was breastfed and the 39% of them was still hospitalized in intensive care unit.

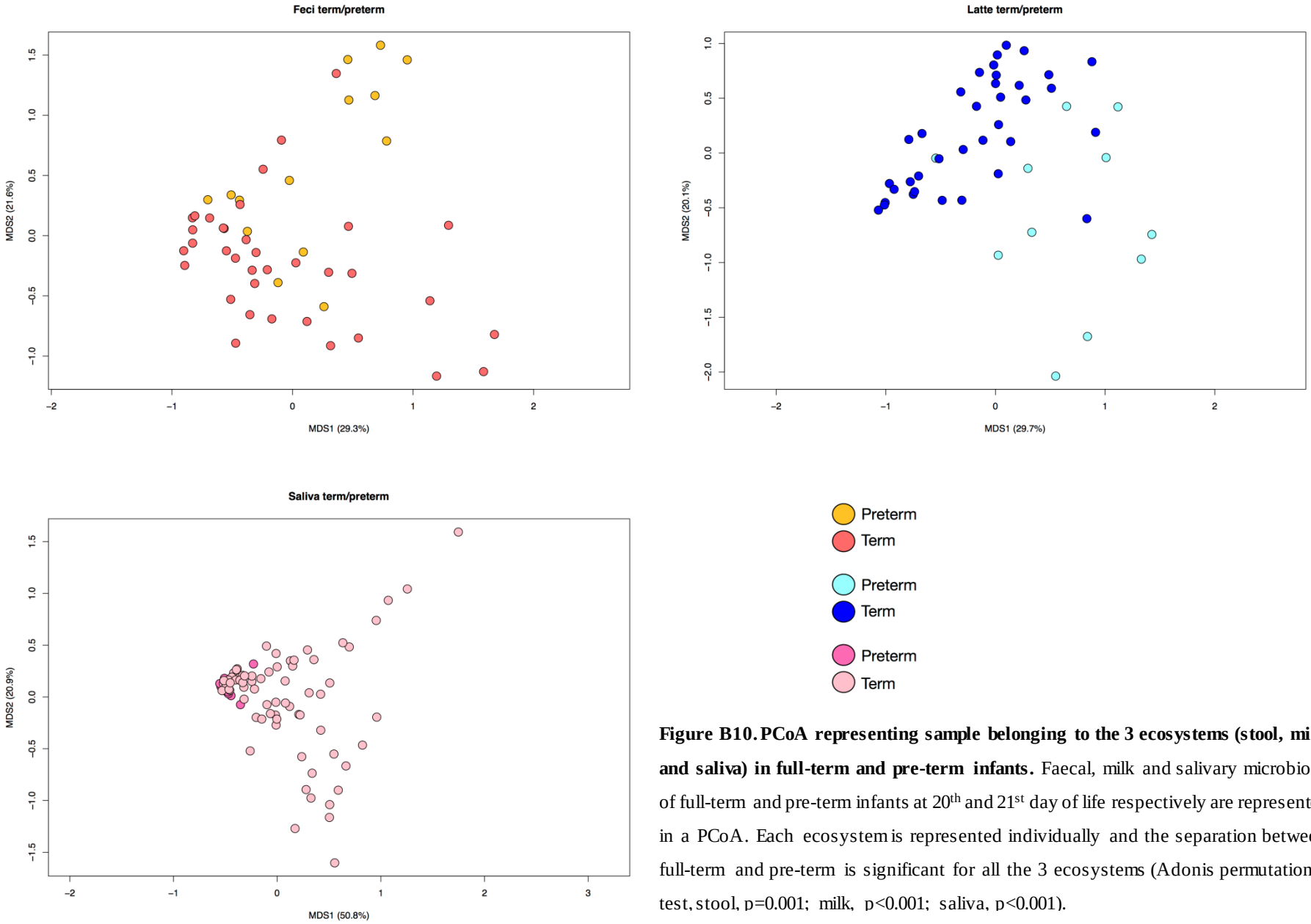
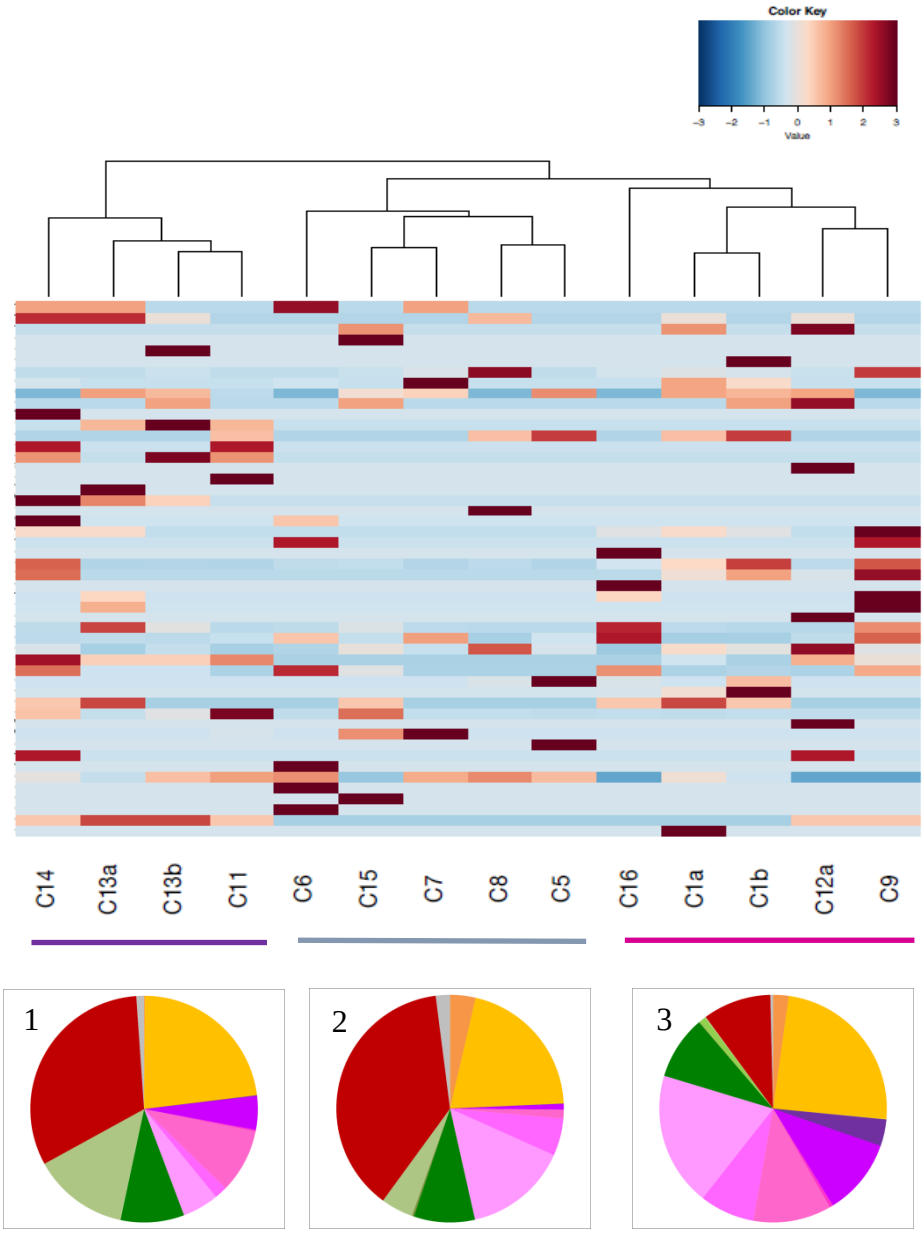


Figure B10. PCoA representing sample belonging to the 3 ecosystems (stool, milk and saliva) in full-term and pre-term infants. Faecal, milk and salivary microbiota of full-term and pre-term infants at 20th and 21st day of life respectively are represented in a PCoA. Each ecosystem is represented individually and the separation between full-term and pre-term is significant for all the 3 ecosystems (Adonis permutational test, stool, p=0.001; milk, p<0.001; saliva, p<0.001).

Because of the high degree of inter-individual variability of faecal samples collected at 21st day of life, it was impossible to mediate the relative abundance composition to compare then the GM structure between full-term and pre-term ones. Nevertheless, a heatmap based on microbial composition described at family level of all the stool samples collected at day 21st of life of pre-term infants was built in order to understand if it could be possible to divide the samples in group according to the microbiological structure. The faecal samples clustered to form 3 groups, and their separation, measured with Adonis permutational test, was significant (p=0.02).



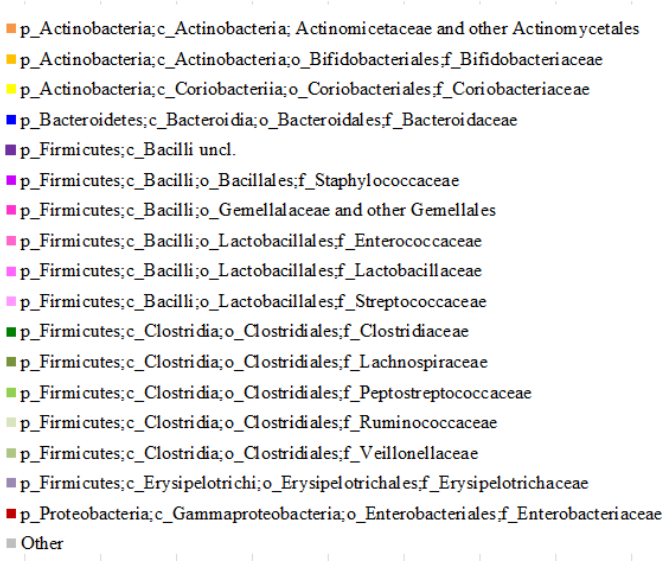


Figure B11. Heat map based on family composition of faecal samples collected at day 21st of life. Heat map shows the relative abundance of GM composition at family level. Hierarchical clustering was performed using the Pearson distance measure and Ward linkage method. The formation of 3 groups is observed and the separation between them is significant (Adonis permutational test, p=0.02). The mean GM structure at family level for each group is represented with a pie chart.

When the microbial composition was mediated between samples belonging to the same group, we noticed that group 1 was characterized by the higher relative abundance of Veillonellaceae, a lactate-fermenting microorganism commonly found in the human intestine. Group 2 was characterized by higher relative abundance of Enterobacteriaceae (37.9%) compared to the other 2 and group 3 by the highest amount of Streptococcaceae, Enterococcaceae and Staphylococcaceae.

The analysis of these data suggests that stool microbial composition differs between full-term and pre-term infants, while the overall composition in terms of dominant families of infant's saliva and mother's milk is comparable. Because of the different drug therapies and feeding conditions the infants were exposed to during their stay in the hospital, it is tempting to speculate that the GM assembly is influenced more by the clinical history and environmental bacteria than the mutual relationship with the mother and the bacteria she provides through latching.

Conclusion and further perspectives

The assembly dynamics of the infant's gut ecosystem is a topic of tremendous interest for human microbiology and immunology (Lynch and Pedersen, 2016). Indeed, the existence of a crucial window of time in which the microbiota contributes to the education of the infant's immune system has been demonstrated (Arrieta *et al.*, 2015; Honda and Littman, 2016). Moreover, if the infant is born before the normal pregnancy term (<37 weeks), the organs immaturity and the pharmacological treatment pave the way to the onset of complications, such as respiratory faint, necrotising enterocolitis and sepsis. The GM has been demonstrated to play a role in the outbreak of some of these complications and for this reason the description of the colonization dynamics of the infant gut during the very first days of life appears to be fundamental.

In the light of these considerations, we characterized the microbial communities of infants' stool and saliva, together with mother breast milk, in two cohorts of babies, one constituted by 36 healthy breast-fed infants born full-term and the other constituted by 21 infants born moderate to late pre-term receiving different types of feeding (mother breast milk human breast milk from donor and formula). The first cohort was sampled at 20th day of life, whilst the second one was sampled longitudinally from birth to 30th day of life.

Our study led on full-term cohort, demonstrate that the microbial ecosystems belonging to other body sites, both of the mother (breast milk) and of the infant itself (saliva) could be involved in the assembly of the gut microbiota. This cohort was very homogenous, and the microbial structures of the ecosystems was comparable between subjects. On the other hand, the infants belonging to the moderately to late pre-term cohort had a very personal clinical history and each baby was a story of its own. The data obtained point out the high inter-individual variability of the ecosystems observed (stools, saliva and milk). When we focused in particular on stools microbial composition, we realize that, as proposed by Zaneveld and colleagues in the "Anna Karenina principle" (2017), the microbiomes of healthy hosts may all look similar, while microbiomes of unhealthy hosts may end up looking very different from one another. Since a dysbiotic GM could lead to the onset of complications linked to the gastrointestinal tract, it is particularly important to have a healthy-like GM structure to have a confrontation with. This may be useful in the frame of setting up an intervention strategy aimed to the restore of a healthy-like GM structure. This inter-individual diversity makes things extremely complicated, but nevertheless, probiotics administration has become a well-founded successful therapy in recent years.

Further works are intended to characterize another category of pre-term infants, namely the extremely pre-term one (<28 weeks). We are recruiting 26 extremely pre-term infants and we have programmed to collect stool samples from the infants and milk sample from the mother at birth and days 1, 4, 7, 30 after birth and at demission from hospital. A very intriguingly feature is that all the samples belonging to the 3 cohorts where were collected in the same hospital. It's therefore highly definable if there's an environmental contamination or a nosocomial infection that could affects the samples

Isolation and characterization of probiotic bacteria from human breast milk

Within the National Technological Cluster Project PROS.IT (Promotion of consumer's health: nutritional valorisation of Italian agrofood traditional products), a research activity directed to the isolation of probiotic bacteria from human breast milk was carried out (*OR 2 - Isolation and characterization of probiotics. Activity 3.1 - Study of new probiotic formulations containing bacteria isolated from human breast milk*). Isolation of bifidobacteria was conducted starting from human breast milk samples collected for the full-term infants' study (Part 1, *The bacterial ecosystem of mother's milk and infant's mouth and gut*). An aliquot of the samples was seeded on 4 different cultural media, before the samples were placed at -80°C until further analysis. The media used were: MRS (De Man, Rogosa and Sharpe) + 0.05% of cysteine, Wilkins-Chalgren, Agar Tos Propionate + mupirocine and RB.

Two strains of bifidobacteria (*Bifidobacterium scardovi* and *Bifidobacterium breve*) and 2 of lactobacillus (*Lactobacillus gasseri* and *Lactobacillus fermentum*) have been isolated using MRS + 0.05% of cysteine, grown at 37°C in an anaerobic chamber containing a gas mix of 90% N_2, 5% CO2 and 5% H_2 for 24 hours. For all microorganisms, resistance to simulated gastric fluid has been tested, according to the methodology described by Fernandez and colleagues (2003). Briefly, a cellular sospension containing 1 x 10^{10} cells were added to 50 ml of simulated gastric fluid (125 mM NaCl, 7 mM KCl, 45 mM $NaHCO_3$, 3 g/L pepsin, pH 2) and incubated in anaerobic condition on a stirring support, in order to simulate peristaltic movements. The survival rate of the microorganisms was determined throughout plate counting of aliquots taken at 30, 60, 90 e 180 minutes of incubation. Plates were incubated in anaerobic condition at 37°C for 24 hours. The test was repeated with another simulated gastric fluid solution at pH 3. Moreover, MIC threshold was tested as required by EFSA, in order to determine if these microorganisms were antibiotic resistant. Finally, the complete sequence of the 16s rDNA was obtained using Sanger sequencing.

Once verified that these bacteria answer to all the features required by WHO in order to be defined "probiotic", they were delivered to Granarolo Company (Bologna) for the developing of a probiotic milk product. This probiotic milk was used in a human trial (*OR 5-In vivo validation of beneficial effects of prototypes of functional foods. Activity-In vivo validation of beneficial effects of functional foods enriched with probiotics*) and was administered to women aged 55-75 years suffering from insomnia for 12 weeks.

References

Aagaard, K., Ma, J., Antony, K.M., *et al*. (2017). The placenta harbors a unique microbiome. Sci Transl. Med., 21:237ra65.

Agostoni, C., Braegger, C., Decsi, T., *et al*. (2009). Breast-feeding: A commentary by the ESPGHAN Committee on Nutrition. J Pediatr. Gastroenterol. Nutr., 49:112-125.

Arboleya S., Binetti, A., Salazar, N. *et al*. (2011). Establishment and development of intestinal microbiota in pre-term neonates. FEMS Microbiol. Ecol., 79:763-772.

Arboleya S., Martinez-Camblor P., Solis G. *et al*. (2017). Intestinal microbiota and weight-gain in pre-term neonates. Front. microbiol., 8:183.

Arrieta, M.C., Stiemsma, L.T., Amenyogbe, N., *et al*. (2014). The intestinal microbiome in early life: health and disease. Front. Immunol. 5, 427.

Arrieta, M.C., Stiemsma, L.T., Dimitriu, P.A., *et al*. (2015). Early infancy microbial and metabolic alterations affect risk of childhood asthma. Sci. Transl. Med. 7: 307ra152.

Bäckhed, F., Ding, H., Wang, T., *et al*. (2004). The gut microbiota as an environmental factor that regulates fat storage. PNAS USA, 2:15718-15723.

Bäckhed, F., Roswall, J., Peng, Y., *et al*. (2015). Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. Cell Host Microbe, 17: 690-703.

Belkaid, Y., and Segre, J.A. (2014). Dialogue between skin microbiota and immunity. Science 346, 954-9. doi: 10.1126/science.1260144

Berrington, J.E., Stewart, C.J., Embleton, N.D., *et al*. (2013). Gut microbiota in pre-term infants: assessment and relevance to health and disease. Arch. Dis. Chil. Fetal. Neonatal. Ed., 9: F286-90.

Biagi, E., Franceschi, C., Rampelli, S. *et al*. (2016). Gut Microbiota and Extreme Longevity. Curr Biol, 26:1480-1485.

Biagi, E., Quercia, S., Aceti A., *et al*. (2017). The bacterial ecosystem of mother's milk and infant's mouth and gut. Front microbiol, 8:1214.

Bik, E.M., Eckburg, P.B., Gill, S.R. *et al*. (2006). Molecular analysis of the bacterial microbiota in the human stomach. PNAS, 17:103:732-737

Bokulich N.A., Chung J., Battaglia T., *et al*. (2016). Antibiotics, birth mode, and diet shape microbiome maturation during early life. Sci. Transl. Med., 8:343ra82.

Bottacini, F., Ventura, M., van Sinderen, D., *et al*. (2014). Diversity, ecology and intestinal function of bifidobacteria. Microb. Cell. Fact., 13: S4.

Brugman, S., Perdijk, O., van Neerven, R.J., *et al*. (2015). Mucosal Immune Development in Early Life: Setting the Stage. Arch. Immunol. Ther. Exp. (Warsz), 63: 251-268.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., *et al*. (2010). QIIME allows analysis of high-throughput community sequencing data. Nat Methods, 7:335-336.

Chu, D.M., Ma, J., Prince, A.L. *et al*. (2016). Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery. Nat. Med., 23: 314-326.

Cong, X., Judge, M., Xu, W. *et al*. (2017). Influence of feeding type on gut microbiome development in hospitalized pre-term infants. Nurs. Res., 66: 123-133.

Costello, E.K., Lauber, C.L., Hamady, M., *et al*. (2009). Bacterial community variation in human body habitats across space and time. Science, 326:1694-1697.

Costello, E.K., Stagaman, K., Dethlefsen, L., *et al*. (2012). The application of ecological theory toward an understanding of the human microbiome. Science 336, 1255-1262.

Costello, E.K., Carlisle, E.M., Bik, E.M., *et al*. (2013). Microbiome assembly across multiple body sites in low-birthweight infants. MBio, 4: e00782-13.

Davé, V., Street, K., Francis, S., *et al*. (2016). Bacterial microbiome of breast milk and child saliva from low-income Mexican-American women and children. Pediatr. Res., 79: 846-854

Ding, T., and Schloss, P.D. (2014). Dynamics and associations of microbial community types across the human body. Nature 509, 357-360.

Dominguez-Bello, M.G., Costello, E.K., Contreras, M., *et al*. (2010) Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. PNAS USA, 29:11971-11975.

Duranti, S., Lugli, G.A., Mancabelli, L. *et al*. (2017). Maternal inheritance of bifidobacterial communities and biphidophages in infants through vertical transmission. Microbiome, 5: 66.

Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 26: 2460-2461.

Fitzstevens, J.L., Smith, K.C., Hagadorn, J.I., *et al*. (2016). Systematic review of the human milk microbiota. Nutr. Clin. Pract., 32: 354-364.

Forsgren M., Isolauri E., Salminen S. *et al*. (2017). Late pre-term birth has direct and indirect effects on infant gut microbiota development during the first six months of life. Acta Paediatr., 106: 1103-1109.

Gibson M.K., Wang B., Ahamadi S. *et al*. (2016). Developmental dynamics of the pre-term infant gut microbiota and antibiotic resistome. Nat Microbiol, 1: 16024.

Goldmann, D.A., Leclair, J. and Macone, A. (1978). Bacterial colonization of neonates admitted to an intensive care environment. J. Pediatr. 93: 288–293.

Goodrich, J.K., Di Rienzi, S.C., Poole, A.C., *et al*., (2016). Conducting a microbiome study. Cell, 158: 250-262.

Gregory, K.E., Samuel, B.S., Houghteling, P. *et al*. (2016). Influence of maternal breast milk ingestion on acquisition of the intestinal microbiome in pre-term infants. Microbiome, 4: 68.

Grice, E.A., and Segre, J.A. (2011). The skin microbiome. Nat. Rev. Microbiol. 9: 244-253

Gritz, E.C., and Bhandari, V. (2015). The human neonatal gut microbiome: a brief review. Front. Pediatr., 3: 17.

Hendricks-Muñoz, K.D., Xu, J., Parikh, H.I., *et al*. (2015). Skin-to-skin care and the development of the pre-term infant oral microbiome. Am. J. Perinatol., 32: 1205-1216

Honda, K. and Littman, D.R. (2016). The microbiota in adaptive immune homeostasis and disease. Nature 535, 75-84.

Jain, N. and Walker, W.A. (2015). Diet and host-microbial crosstalk in postnatal intestinal immune homeostasis. Nat. Rev. Gastroenterol. Hepatol., 12: 14-25.

Jost, T., Lacroix, C., Braegger, C. *et al*. (2013). Assessment of bacterial diversity in breast milk using culture-dependent and culture-independent approaches. Br. J. Nutr., 110: 1253-62.

Jost, T., Lacroix, C., Braegger, C.P. *et al*. (2012). New insights in gut microbiota establishment in healthy breast fed neonates. PLoS One, 7: e44595.

Jost, T., Lacroix, C., Braegger, C.P., *et al*. (2014). Vertical mother-neonate transfer of maternal gut bacteria via breastfeeding. Environ. Microbiol., 16: 2891-2904.

Kim, Y.J., Choi, Y.S., Baek, K.J., et al. (2016). Mucosal and salivary microbiota associated with recurrent aphthous stomatitis. BMC Microbiol., 1:1:57.

Klindworth A., Pruesse E., Schweer T. et al. (2013). Evaluation of general 16S ribosomal RNA genePCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Research, 41:e1.

Li, K., Bihan, M., and Methé, B.A. (2013). Analyses of the stability and core taxonomic memberships of the human microbiome. PLoS One, 8: e63139.

Lif Holgerson, P., Öhmanm C., Rönnlund, A. *et al*. (2015). Maturation of oral microbiota in children with or without dental caries. PLoS One, 10: e0128534.

Lin, R., Liu, W., Piao, M., *et al*. (2017). A review of the relationship between the gut microbiota and amino acid metabolism. Amino Acids, 49: 2083-2090.

Lozupone, C.A., Hamady, M., Kelley, S.T. *et al*. (2007). Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities. Appl. Environ. Microbiol., 73, 1576–1585.

Lynch, S.V., and Pedersen, O. (2016). The human intestinal microbiome in health and disease. N. Engl. J. Med., 375: 2369-2379.

Masella A.P., Bartram A.K., Truszkowski J.M. *et al*. (2012). PANDAseq: paired-end assembler for illumina sequences. BMC Bioinformatics, 13: 31.

Meadow, J.F., Altrichter, A.E., Bateman, A.C., *et al*. (2015). Humans differ in their personal microbial cloud. PeerJ 3, e1258. doi: 10.7717/peerj.1258

Meinzen-Derr, J., Morrow, A.L., Hornung, R.W., *et al*. (2009). Epidemiology of necrotizing enterocolitis temporal clustering in two neonatology practices. J. Pediatr, 154: 656-661.

Mizuno, K., and Ueda, A. (2003). The maturation and coordination of sucking, swallowing, and respiration in pre-term infants. J. Pediatr., 142: 36–40

Mueller, N.T., Bakacs, E., Combellick, J., *et al*. (2015). The infant microbiome development: mom matters. Trends Mol. Med., 21: 109-117

Pammi M., Cope J., Tarr P. *et al*. (2017). Intestinal dysbiosis in pre-term infants preceding necrotizing enterocolitis: a systematic review and meta-analysis. Microbiome, 5: 31.

Pannaraj P.S., Li F., Cerini C. *et al*. (2017). Association between breast milk bacterial communities and establishment and development of the infant gut microbiome. JAMA Pediatrics, 171: 647-654.

Penders J., Gerhold K., Thijs C. *et al*. (2014). New insight into the hygiene hypothesis in allergic diseases. Gut microbes, 5: 239-244.

Qin, J., Li, R., Raes, J., *et al*. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. Nature, 464: 59–65.

Rajilic-Stojanovic, M., Heilig, H., Molenaar, D., *et al*. (2009). Development and application of the human intestinal tract chip, a phylogenetic microarray: analysis of the universally conserved phylotypes in the abundant microbiota of young and elderly adults. Environ. Microbiol., 11: 1736-51.

Riezzo, G., Indrio, F., Montagna, O., *et al*., (2000). Gastric electrical activity and gastric emptying in term and pre-term newborns. Neurogastroenterol. Motil., 12: 223–229.

Rodríguez, J.M. (2014). The origin of human milk bacteria: is there a bacterial entero-mammary pathway during late pregnancy and lactation? Adv. Nutr. 5, 779-784.

Round, J. L., and Mazmanian, S. K. (2009). The gut microbiota shapes intestinal immune response during health and disease. Nat Rev Immunol, 9: 313-323.

Ruiz, L., Moles, L., Gueimonde, M, et al. (2016). Perinatal Microbiomes' Influence on Pre-term Birth and Pre-terms' Health. J. Pediatr. Gastroenterol. Nutr. 63: e193–e203.

Russell, W. R., Hoyles, L., Flint, H. J., *et al*. (2013). Colonic bacterial metabolites and human health. Curr. Opin. Microbiol., 16, 246–254.

Sampaio-Maia, B., and Monteiro-Silva, F. (2014). Acquisition and maturation of oral microbiome throughout childhood: An update. Dent. Res. J. (Isfahan), 11: 291-301.

Schloss, P.D., and Westcott, S.L. (2011). Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. Appl. Environ. Microbiol., 77: 3219-26.

Shapiro-Mendoza, C. K., and Lackritz, E. M. (2012). Epidemiology of late and moderate pre-term birth. Semin. Fetal Neonatal Med., 17: 120–125.

Sherman MP, Sherman J, Arcinue R, *et al*. (2016). Randomized Control Trial of Human Recombinant Lactoferrin: A Substudy Reveals Effects on the Fecal Microbiome of Very Low Birth Weight Infants. J. Pediatr, 173: S37-S42.

Song, S.J., Lauber, C., Costello, E.K., *et al*. (2013). Cohabiting family members share microbiota with one another and with their dogs. Elife, 2: e00458.

Stahringer, S.S., Clemente, J.C., Corley, R.P., *et al*. (2012). Nurture trumps nature in a longitudinal survey of salivary bacterial communities in twins from early adolescence to early adulthood. Genome Res., 22: 2146-2152.

Staiano, A., Boccia, G., Salvia, G., *et al*. (2007). Development of esophageal peristalsis in pre-term and term neonates. Gastroenterology 132: 1718–25.

Stewart C.J., Embleton N.D., Marrs E.C.L. *et al*. (2016). Temporal bacterial and metabolic development of the premature gut reveals specific signatures in health and disease. Microbiome, 4: 67.

Stewart C.J., Embleton N.D., Marrs E.C.L. *et al*. (2017). Longitudinal development of the gut microbiome and metabolome in pre-term neonates with late onset sepsis and healthy controls. Microbiome, 5: 75.

Tannock GW. 1999. Analysis of the intestinal microflora: a reinassance. Antonie van Leeuwenhoek 76, 265.

Wade, W.G. (2013). The oral microbiome in health and disease. Pharmacol. Res., 69: 137-143

Walker, A.W., Martin, J.C., Scott, P., *et al*. (2015). 16S rRNA gene-based profiling of the human infant gut microbiota is strongly influenced by sample processing and PCR primer choice. Microbiome, 3: 26.

Wampach L., Heintz-Buschart A., Hogan A. *et al*. (2017). Colonization and succession within the human gut microbiome by archea, bacteria and microeukaryotes during the first year of life. Front microbiol, 8: 738.

Wheeler, M.L., Limon, J.J., Bar, A.S., *et al*., (2016). Immunological Consequences of Intestinal Fungal Dysbiosis. Cell Host Microbe, 8: 865-873.

Yu, Z. and Morrison, M. (2004). Improved extraction of PCR-quality community DNA from digesta and fecal samples. Biotechniques, 36: 808-812.

Zaura, E., Nicu, E.A., Krom, B.P. *et al*. (2014). Acquiring and maintaining a normal oral microbiome: current perspective. Front. Cell. Infect. Microbiol., 4: 85.

Zoetendal, EG, Rajilic-Stojanovic, M, and de Vos, WM. (2008). High-throughput diversity and functionality analysis of the gastrointestinal tract microbiota. Gut, 57: 1605-1615.

Zaneveld, J.R, McMinds, R. and Thurber, R.V. (2017). Stress and stability: applying the Anna Karenina principle to animal microbiomes. Nat. Microbiol., 2: 17121.