Oral microbiome development during childhood: an ecological succession influenced by postnatal factors and associated with tooth decay

Short title: Oral microbiome development during childhood

Authors: Majda Dzidic 1,2,3, Maria Carmen Collado 2, Thomas Abrahamsson 4, Alejandro Artacho 1, Malin Stensson 5, Maria C Jenmalm 3 and Alex Mira 1*

Affiliations:
1. Department of Health and Genomics, Center for Advanced Research in Public Health, CSISP-FISABIO, Valencia, Spain
2. Institute of Agrochemistry and Food Technology (IATA-CSIC), Department of Biotechnology, Unit of Lactic Acid Bacteria and Probiotics, Valencia, Spain
3. Department of Clinical and Experimental Medicine, Division of Autoimmunity and Immune Regulation, Linköping University, Linköping, Sweden
4. Department of Clinical and Experimental Medicine, Division of Pediatrics, Linköping University, Linköping, Sweden
5. Centre for Oral Health, School of Health and Welfare, Jönköping University, Sweden

Correspondence:
*To whom correspondence may be addressed.
Alex Mira
Address: Avenida de Cataluña, 21 46020 Valencia
Tel.: 961 92 59 09
Fax: 961 92 57 03
Email: mira_ale@gva.es (A.M.)

Conflict of interests

The authors declare that they have no competing interests.
Subject category: Microbe-Microbe and microbe-host interactions

Funding
A.M.: Spanish Ministry of Economy and Competitiveness (grant no. BIO2015-68711-R). M. S.: The Research Council for the South-East Sweden (grant no: 79001). M. C. J.: The Swedish Research Council (2016-01698); the Swedish Heart and Lung Foundation (20140321); the Medical Research Council of Southeast Sweden (FORSS-573471); the Cancer and Allergy Foundation. M. C. C.: European Research Council (ERC-starting grant 639226).
ABSTRACT

Information on how the oral microbiome develops during early childhood and how external factors influence this ecological process is scarce. We used high-throughput sequencing to characterise bacterial composition in saliva samples collected at 3, 6, 12, 24 months and 7 years of age in 90 longitudinally followed children, for whom clinical, dietary and health data were collected. Bacterial composition patterns changed through time, starting with “early colonizers”, including Streptococcus and Veillonella; other bacterial genera such as Neisseria settled after one or two years of age. Dental caries development was associated with diverging microbial composition through time. Streptococcus cristatus appeared to be associated with increased risk of developing tooth decay and its role as potential biomarker of the disease should be studied with species-specific probes. Infants born by C-section had initially skewed bacterial content compared to vaginally delivered infants, but this was recovered with age. Shorter breastfeeding habits and antibiotic treatment during the first 2 years of age were associated with a distinct bacterial composition at later age. The findings presented describe oral microbiota development as an ecological succession where altered colonization pattern during the first year of life may have long-term consequences for child’s oral and systemic health.

INTRODUCTION

The development and structure of the neonatal microbiome have been partially elucidated, with a main focus on the microbial population inhabiting the lower intestinal tract, while information about the oral cavity colonization following delivery is still limited (Gomez & Nelson, 2017). As yet, no published longitudinal studies have characterized oral microbiota development during infancy and childhood with culture independent next generation sequencing methodologies, particularly in association with tooth decay.

It is believed that by production and excretion of metabolic products of pioneer colonizers (including facultative anaerobes Streptococcus and Actinomyces), acquired at birth and the following hours, the environment can be altered, thus benefiting and selecting the growth of other species (including more strictly anaerobic genera like Veillonella and Fusobacteria) (Gomez & Nelson, 2017; Sampaio-Maia & Monteiro-Silva, 2014). As the baby grows, microbial communities evolve and increase in microbial diversity (Cephas et al., 2011; Lif Holgerson et al., 2015). During this period the oral microbiota is characterized by high variability and current knowledge indicates that it reaches adult-like stability around two years of age (Gomez & Nelson, 2017).
Most evidence available today shows that the early oral environment is strongly shaped by the mother (Flores et al., 2014; Sampaio-Maia & Monteiro-Silva, 2014; Teng et al., 2015) and maternal oral microbiota has been proposed to colonize the placenta (Aagaard et al., 2014) where it could influence fetal immune tolerance towards the mother’s microbiome (Zaura et al., 2014). Further transition into a more mature and complex microbial ecosystem is mainly influenced by the external environment as well as vertical transmission from the parents (Flores et al., 2014; Hesselmar et al., 2013; Sampaio-Maia & Monteiro-Silva, 2014; Song et al., 2013). An essential question is to identify which factors and at what time point they can influence the progression of microbial colonization. Previous studies of the lower gastrointestinal tract microbiota have reported that the gut microbiota of infants delivered by caesarean section (C-section) was mainly colonized by skin bacteria, had lower numbers of \textit{Bifidobacterium} and \textit{Bacteroides} species and were more often colonized with \textit{Clostridium difficile} in comparison to vaginally born infants (Jakobsson et al., 2014; Penders et al., 2006). However, research regarding the influence of delivery mode on the early oral microbiota development, by using next generation sequencing on longitudinal samples, has not yet been reported.

Breast milk has long been considered a superior food for infants, increasing resistance to infections, providing nutrition and being a source of bacteria (10^6 bacterial cells/ml), who serve as inoculum for the newborn (Boix-Amorós et al., 2016; Fernández et al., 2013; Fitzstevens et al., 2016; Rodriguez, 2014). The genus \textit{Streptococcus} is one of the dominant bacterial groups found in human milk (Boix-Amorós et al., 2016; Fitzstevens et al., 2016) and various species, including \textit{Streptococcus salivarius}, are frequently found in the infant oral cavity (Carlsson et al., 1970). The metabolic products derived from \textit{Streptococcus} species from the dietary oligosaccharides in breast milk might pave the way for the establishment of other microorganisms in the oral cavity, thus influencing attachment and growth of selected bacteria (Aimutis, 2004; Danielsson Niemi et al., 2009; Gomez & Nelson, 2017; Sampaio-Maia & Monteiro-Silva, 2014; Sheedy et al., 2009; Wernersson et al., 2006). However, the longitudinal impact of these initial colonizers on the oral ecosystem and the influence of breastfeeding habits on children’s oral and systemic health are widely unknown and deserve to be investigated.

Knowledge about the effect of other external factors like antibiotic use, especially at an early age, on subsequent microbiome development is also scarce. In children, long-term alterations of the gut microbiome as a consequence of early antibiotic administration have been described and proposed to have negative effects for systemic health, including obesity and allergy (Ajslev et al., 2011;
Reynolds & Finlay, 2017). However, the long-term effect of antibiotic use for children’s oral microbiota is currently unknown.

An important consequence of oral microbiome development for health is the protection against tooth decay (dental caries), considered among the most prevalent diseases worldwide (Petersen, 2003). Tooth decay is caused by an interaction between acidogenic bacteria, a carbohydrate substrate and host susceptibility, leading to bacterial dysbiosis and demineralization of tooth tissue (Lif Holgerson et al., 2015; Selwitz et al., 2007). The acid-tolerant bacterial species *Streptococcus mutans* is recognized to be an important pathogen in dental caries, (E Kanasi et al., 2010; Tanner et al., 2011) and its early presence in edentulous children (from 3 months of age), is suggesting that the soft tissue may play a role as a reservoir for oral pathogenic microorganisms (Cephas et al., 2011; Nelun Barfod et al., 2011). Given that early colonization with cariogenic microorganisms has been associated with higher caries incidence (E Kanasi et al., 2010), microbiological studies in longitudinal samples through early childhood may reveal those bacteria increasing caries risk that could be used as early diagnostic biomarkers. This could also provide important information for active and passive immunization strategies against oral diseases (Abiko, 2000). Moreover, an unhealthy oral microbiome can have important effects beyond the oral cavity, including elevated cardiovascular risk (Mathews et al., 2016; Scannapieco et al., 2003). For instance, *in vitro* studies have demonstrated the ability of periodontal bacteria to increase the probability of thrombus formation, which could lead to ischemic cardiovascular events (Demmer & Desvarieux, 2006; Fong, 2000). Therefore, it is of interest to understand the colonization patterns of oral commensals during childhood and the potential benign effect of oral bacteria in preventing oral and systemic diseases, including microorganisms which have been associated with health conditions (Huang et al., 2015; López-López et al., 2017).

A more detailed understanding of oral microbial communities development in health and disease fundamental and the use of high-throughput sequencing techniques now allow exploring microbial composition and diversity in low volume oral samples to an unprecedented level of detail (Nyvad et al., 2013), in comparison with culturing or early molecular methodologies. In this study, we aimed to address the temporal evolution and maturation of the oral microbial ecosystem during infancy and childhood and its relation to delivery mode, breastfeeding habits, antibiotic use and dental caries status, in longitudinally collected oral samples in 90 children followed from birth to seven years of age.
METHODS

Sample collection and study design

The infants included in the study were part of a larger randomized double-blind trial in Sweden between 2001 and 2003 evaluating the potential allergy prevention effect of probiotic *Lactobacillus reuteri* ATCC 55730 until 2 and 7 years of age (Abrahamsson et al., 2007; Thomas R. Abrahamsson et al., 2013). Among the 188 infants completing the original study, longitudinal salivary samples were collected in 90 children. The participants were instructed not to eat or drink for two hours preceding the sampling. Non-stimulated saliva samples at 3, 6, 12 and 24 months of age were collected from the buccal cavity, using a hand pump (Nalgene #6131, ThermoFisher, Stockholm, Sweden) connected to a thin plastic tube and immediately frozen and kept at −80°C. At 7 years of age, paraffin-stimulated whole saliva was collected (≈3 ml) in a sterile test tube and immediately frozen at −80°C. By 9 years of age, the children were examined at public dental clinics at which the children received their regular dental care (Stensson et al., 2014), and the caries status was evaluated. The oral examination included radiographs and the registration of manifest and initial caries lesions in the primary dentition according to Koch et al. and Alm et al. (Alm et al., 2007; Koch, 1967).

Possible confounders, such as mode of delivery, breastfeeding habits (exclusive or partial breast feeding), infant health and antibiotics use during the first two years of age were obtained from medical records and semi-structured questionnaires (see Table 1) (Stensson et al., 2014). 91% and 80% of all children included were exclusively breast-fed up to 1 and 3 months of age, respectively, while 97% were partially breastfed at 3 months of age. No infant received antibiotics before 1 month of age while 2% took antibiotics during the first 3 months of life.

The studies were approved by the Regional Ethics Committee for Human Research in Linköping, Sweden (Dnr 99323, M122-31 and M171-07, respectively). An informed consent was obtained from both parents before inclusion in the study. Written informed consent was also given by the parents or guardians before the dental examination.

DNA extraction

250 ul of each saliva sample were centrifuged at 15000 g for 30 min and the pellet, together with 50 ul of the supernatant, was used for further analysis. DNA was isolated by MagNA Pure LC 2.0 Instrument (1996-2016 Roche Diagnostics, Barcelona, Spain), using MagNA Pure LC DNA Isolation Kit III for Bacteria and Fungi (Roche Diagnostics GmbH, Mannheim, Germany) following...
the manufacturer’s instructions with an additional enzymatic lysis step with lysozyme (20 mg/ml, 37°C, 60 min; Thermomixer comfort, Eppendorf, Hamburg, Germany), lysostaphin (2000 units/mg protein, 37°C, 60 min; Sigma-Aldrich, Madrid, Spain) and mutanolysin (4000 units/mg protein, 37°C, 60 min; Sigma-Aldrich). DNA was resuspended in 100 ul of elution buffer and frozen at -20°C until further analysis.

**16S rRNA gene amplification and sequencing**

Prior to sequencing of 16S rRNA gene, extracted DNA was pre-amplified by using universal bacterial degenerate primers 8F-AGAGTTTGATCMTGGCTCAG and 926R-CCGTCATTGCTTTRAGT, which encompass the hypervariable regions V1-V5 of the gene. This was performed using the high-fidelity AB-Gene DNA polymerase (Thermo Scientific, Waltham, Mass., USA) with an annealing temperature of 52°C and 10 cycles, in order to minimize amplification biases (Sipos et al., 2007). The purification of PCR products was completed using Nucleofast 96 PCR filter plates (Macherey-Nagel, Düren, Germany).

An Illumina amplicon library was performed following the 16S rRNA gene Metagenomic Sequencing Library Preparation Illumina protocol (Part #15044223 Rev. A). The gene-specific primer sequences used in this protocol were selected from Klindworth et al. (Klindworth et al., 2013) and target the 16S rRNA gene V3 and V4 regions, resulting in a single amplicon of approximately 460 bp. Overhang adapter sequences were used together with the primer pair sequences for compatibility with Illumina index and sequencing adapters. After 16S rRNA gene amplification, the DNA was sequenced on a MiSeq Sequencer according to manufacturer’s instructions (Illumina) using the 2x300 bp paired-end protocol. Sequences supporting the conclusions of this article are publicly available at the European Nucleotide Archive (ENA) database with the accession number PRJEB66628.

**Bacterial load and Streptococcus dentisani measurements with quantitative PCR**

Total bacterial load (number of bacterial cells per ml of saliva) and the presence of *Streptococcus dentisani* in saliva samples were measured by quantitative PCR. Amplifications were performed in duplicates on a LightCycler 480 Real-Time PCR System (Roche Technologies) by using annealing temperatures of 60°C and 65°C for total bacterial load and *S. dentisani*, respectively. Each reaction mixture of 10 mL was composed of SYBR Green PCR Master Mix (Roche), 0.5 mL of the specific primer (concentration 10 mmol/L), and 2 mL of DNA template. For *S. dentisani* the forward primer was 5’GTA ACC AAC CGC CCA GAA GG 3’ and the reverse primer 5’CCG CTT TCG GAC TCG ATC A 3’ (Integrated DNA Technologies (IDT); San Diego, California, USA) targeting the
carbamate kinase gene, and for total bacterial density measurements the universal forward and reverse primers were 5’GTG CCA GCM GCC GCG GTA A 3’ and 5’GCG TGG ACT ACC AGG GTA TCT 3’ (IDT), respectively, targeting the bacterial 16S rRNA gene. The obtained Ct values were transformed in bacterial cell numbers by a standard curve calibrated by flow cytometry (Boix-Amorós et al., 2016).

**Bioinformatics and statistics**

Only overlapping paired end reads were used for analysis. A sequence quality assessment was carried out using the PRINSEQ program (Schmieder & Edwards, 2011). Sequences of <250 nucleotides in length were not considered; sequence end-trimming was performed by cutting out nucleotides with a mean quality of <30 in 20-bp windows. Chimeric 16S sequences were filtered out using USEARCH program (Edgar, 2016).

Obtained sequences were taxonomically classified by the RDP-classifier (Wang et al., 2007) where reads were assigned a phylum, class, family and genus and phylogenetic ranks were allocated when scores exceeded an 0.8 confidence threshold. Operational taxonomic units (OTUs) were generated by using CD-HIT OTU picking with 97% of similarity (Li & Godzik, 2006). Human oral microbiome database (HOMID) was used as a reference database for OTU assignment (Chen et al., 2010). For the *Streptococci*-species analyses, sequences were clustered into operational taxonomic units (OTUs) at 100% similarity by BLAST analysis (Altschul et al., 1990) and > 350 bp alignment length, against the RDP database (Cole et al., 2014). A few species appeared to be identical in the sequenced region, namely *Streptococcus infantis*, *S. mitis* and *S. dentisani*, and could not be distinguished from each other.

α–diversity analyses (presented here as Shannon and Chao1 indices), were utilized to estimate samples’ diversity and richness at the 97% OTU level using the R-package Vegan (Oksanen, 2018). Constrained correspondence analysis (CCA, a.k.a. canonical correspondence analysis) is a statistic tool used to emphasize variation, taking advantage of the fact that the factor provided can explain part of the total variability, and bring out strong patterns in a dataset. This analysis was performed by R software ade4 package (Dray S. and Dufour AB., 2007) using the function CCA, which is based on Chi-squared distances. Adonis tests were done with the R library ‘vegan’ (Oksanen, 2018). It performs a permutational multivariate analysis of variance using distance matrices and fitting linear models to them. The test allows modelling the whole compositional variability at once by taking into account different sources of variation as well as interactions between them as it is defined in a linear model.
Linear discriminant analysis effect size (LEfSe), a method for biomarker discovery on the online interface Galaxy (http://huttenhower.sph.harvard.edu) (Segata et al., 2011), was used to detect the taxa, at both genus and OTU level, characterizing the populations of caries-free and caries active children.

Statistical analyses were performed in R version 3.2.2 and GraphPad Prism 6 (GraphPad Software, San Diego, CA, USA, Version 6.1f), where p<0.05 was considered significant. Specific statistical tests (including Mann-Whitney U-test for nonparametric comparisons) are stated in figure legends. When comparing the frequencies of different bacterial taxa between groups (e.g. caries-free and caries-experienced children), the balanced proportions of confounding factors, including breastfeeding length, mode of delivery and antibiotic intake, were checked by Chi-square test, and non-significant differences between the groups were found.
FINDINGS AND DISCUSSION

After quality filtering, 34,794,056 sequences were obtained, with an average of 93,532 ± 3,480 (SEM) sequences per sample.

Bacterial load, richness and diversity through time

Bacterial diversity and richness increased through time, reaching nearly 550 OTUs at 7 years of age with a Shannon diversity index of approximately 2.4 (Fig. 1). The delivery mode and partial breastfeeding habits until 12 months of age did not have an impact on species richness (Fig. 1a-b). However, bacterial diversity appeared to be higher in C-section delivered infants at 12 months of age (Fig. 1a) and at two years of age in children not being breastfed through 12 months of age (Fig. 1b).

Oral development, including the emergence of teeth, was accompanied by a steady increase in diversity and richness of the oral microbiome in this study, especially between 1 and 2 years of age. Interestingly, bacterial diversity at 2 years of age (Fig. 1b), appears to be higher in children which abandoned breastfeeding before 12 months of age. Although this has not been studied before in oral microbiota, a similar trend was observed in gut microbiota analyses where children not being breastfed had higher microbial diversity (Thomas Abrahamsson et al., 2013; Azad et al., 2013; Bäckhed et al., 2015), probably due to earlier introduction of solid food. Our results agree with a scenario in which following delivery, the oral cavity gets exposed to the environment, triggering the initiation of microbial colonization through diet, vertical transmission from parents and horizontal transmission from caregivers and siblings, thus increasing the bacterial diversity (Könönen, 2000; Nelson-Filho et al., 2013).

In order to determine the development of bacterial density through infancy, we measured total bacterial load (Fig. SI1) in saliva samples. Although there were no differences regarding delivery mode (Fig. SI1a) and breastfeeding habits (Fig. SI1b), the density of bacteria increased significantly with age, probably reflecting the influence of environmental interactions and the emergence of teeth. Interestingly, bacterial density at each time point appeared to fall within two groups (low or high), and this bimodal distribution was maintained through time for each individual. This pattern could not be attributed to caries status, allergy development, mode of delivery, feeding habits, antibiotics intake or probiotic administration (data not shown). In the future, it would be interesting to determine whether the physicochemical properties of saliva may influence cell density.
Bacterial composition during infancy

When bacterial composition was analysed for all samples through child development, clear changes emerged through time (Fig. SI2). Streptococci dominated salivary samples at all times. They were particularly high in proportion during the first months of age, and their decrease was accompanied by a rise in other genera. These general patterns were influenced by several perinatal and postnatal factors.

The influence of delivery mode and breastfeeding durations

Bacterial species composition development was influenced by delivery mode and breastfeeding habits (Fig. 2a-b), but not by L. reuteri supplementation during the first year of age (data not shown). The impact of delivery mode was reflected in differences in bacterial composition at 3 and 6 months of age (Fig. 2a, p=0.001, CCA analysis), followed by convergent microbial patterns at later age. Only the genus Haemophilus was found to be significantly more abundant (p=0.047) at 7 years of age in children delivered by C-section (Fig. SI3). Thus, with the exception of this genus, no further colonizers were found to be significantly different between vaginally delivered and C-section delivered infants (Fig. SI3). This could be due to infant delivery mode affecting the direct transmission of initial bacteria from mother to newborn, having a short-term effect. This finding is in line with previous studies (Lif Holgerson et al., 2011) where the Human Oral Microbe Identification Microarray was used, showing that microbial oral colonization in three-month-old infants delivered vaginally and those delivered by C-section was different. Similar findings of an early impact, but also more long-term effects (Hyde & Modi, 2012; Jakobsson et al., 2014), have been reported for the microbiota of the lower gastrointestinal tract (Dominguez-Bello et al., 2010; Penders et al., 2006). When a multivariate analysis was performed including time, breastfeeding length and caries status as confounding factors, the effect of delivery mode on microbiota composition was no longer significant. Given that a significant breastfeeding length-delivery mode interaction was detected (p=0.026), part of the observed differences between children born by vaginal delivery and C-section can be due to the effect of breastfeeding.

The influence of partial compared to no breastfeeding until 12 months of age did appear to have a long-term effect, as evidenced by a divergent oral bacterial composition at 24 months and 7 years of age (Fig. 2b, p=0.002) while bacterial colonization at early age appeared to be similar. This could be due to the fact that the majority of the infants in this cohort were breastfed during their first months of life (see Table 1). A multivariate analysis revealed that the significant effect of breastfeeding on microbiota composition was maintained even after removing the effect of caries...
status, time and mode of delivery as confounding factors (p=0.036). Further work should therefore address the impact of formula feeding on microbiome development as findings presented here suggest that variations in the initial oral microbial communities may result in differences in the bacterial succession patterns that persist over time, analogous to the impact of early disturbance in ecological successions (Amarasekare & Possingham, 2001).

Microbial colonization patterns

Dominant bacterial genera (present at >1%) which inhabited the oral cavity during the first 3-6 months, here called “Early colonizers”, included Streptococcus, Veillonella and Lactobacillus spp. (Fig. 3a). The most frequent bacterium of the oral cavity in the current study was Streptococcus, and children being breastfed until 12 months of age appeared to have higher abundance of this genus at one year of age (p=0.005). This finding is consistent with other reports (Cephas et al., 2011; Luo et al., 2012) and Streptococcus has been found to be one of the dominant bacterial groups in human breast milk (Boix-Amorós et al., 2016; Rodriguez, 2014). Aging of the children was associated with lower levels of Streptococcus, although the decrease tended to be more notable in children abandoning breastfeeding before 12 months of age. This indicated that settlement of this genus is favoured by breast milk, either by direct transmission or by an appropriate nutrient supply (Boix-Amorós et al., 2016; Hunt et al., 2011). Moreover, this pioneer is often found in the oral cavity of the neonate because of its ability to adhere to and colonize the mucosal surface lining (Sampaio-Maia & Monteiro-Silva, 2014). The metabolic products (such as lactic acid) derived from Streptococcus species from the dietary oligosaccharides in breast milk might pave the way for the establishment of other microorganisms in the oral cavity, including bacterial genera like Veillonella (Gomez & Nelson, 2017; Wernersson et al., 2006). Veillonella, here ranging between 2 to 8% of total abundance with significantly higher levels at 7 years in children keeping breastfeeding until 12 months of age (p=0.037), is another bacterial genus commonly encountered in breast milk (Cabrera-Rubio et al., 2012; Jost et al., 2014). This genus requires organic acids as carbon source and therefore its presence is likely favoured by the high levels of lactate derived from lactose fermentation, which this genus will transform to propionate and acetate (Jost et al., 2015). An important lactose fermenter is obviously Lactobacillus, which in the oral cavity might be acquired by the neonate during vaginal delivery, as this genus is highly abundant in vaginal microbiota, (Martin et al., 2012) but also through breastfeeding since breast milk has been proposed to favour the growth of vaginally acquired bacteria (Dominguez-Bello et al., 2010; Jost et al., 2015; Soto et al., 2014). In the current study, no differences in Lactobacillus abundance were found between children being breastfed up to one year of age or not (Fig. 3a) and neither between vaginally delivered and C-section infants (Fig. S13). Among the components of human breast milk,
oligosaccharides are thought to directly influence the gut microbial composition and to enrich bacterial functions associated to carbohydrate consumption and biosynthesis of amino acids and vitamins (Bäckhed et al., 2015; Marcobal et al., 2010) and a similar process may be taking place in the oral cavity. Early commensals of the oral cavity are likely having an ecological advantage over those arriving later and may promote the change of the environment through the production and excretion of products of their metabolism, thus benefitting the growth of further oral bacterial communities. This process of microbial succession and increasing diversity, promoted by breastfeeding, could lead to subsequent formation of complex and steadier microbial communities, as proposed for gut microbiota (Sprockett et al., 2018).

Bacterial genera Gemella, Granulicatella, Haemophilus and Rothia, here defined as “constant colonizers” (Fig. 3b), were present already at 3 and 6 months of age with >1% of abundance, and their abundance increased with time. Gemella and Granulicatella are considered as common dental plaque inhabitants (Aas et al., 2005) and were found to increase in abundance through age, ranging from 5-10% and 2-8%, respectively. It is likely that the initiation of teeth eruption, starting around 6-8 months postnatally, creates new ecological niches in the oral cavity, giving rise to new adhesion surfaces, thus favouring their further colonization.

A third set of microorganisms were “late colonizers” and included Actinomyces, Porphyromonas, Abiotrophia and Neisseria, which became dominant in the oral cavity at a later stage, approximately after the first year of life (Fig. 3c). Thus, the data suggest that the acquisition or dominance of each bacteria may occur optimally only at certain ages. Breastfeeding until 12 months of age was associated with significantly lower levels of Actinomyces (p=0.044) at 7 years of age and Porphyromonas (p=0.049) and Neisseria (p=0.028) at 12 months and 24 months of age, respectively. Porphyromonas, more specifically Porphyromonas gingivalis, is a gram-negative oral anaerobe involved in the pathogenesis of periodontitis, an inflammatory disease that destroys the tooth tissue and may lead to tooth loss (Mysak et al., 2014). The results are indicating that children being breastfed by 12 months of age, as compared with children no longer breastfed, have significantly lower abundance of this genus at one year of age. However, species-level taxonomic analysis revealed that 100% of Porphyromonads sequences correspond to Porphyromonas catoniae during the first 12 months of age. At 2 years, P. gingivalis appeared at 9% of the total, whereas P. catoniae accounted for 91% of the sequence reads. At 7 years of age, the proportions were 86.5% for P. catoniae and 13.4% P. gingivalis. Thus, an association between reduced breastfeeding length and risk of gum disease is uncertain. Neisseria, a common bacterial community member of the healthy human mouth (Bik et al., 2010), was found to be more abundant in children not being
breastfed until 12 months of age, in line with previous research where species belonging to this genus were found more frequently in children being formula-fed (Holgerson et al., 2013). Thus, breast milk had a long-term effect on oral microbiota composition, but this altered microbiota could not always be linked to healthy or disease-associated communities, and further work should study the long-term consequences for the child’s oral and systemic health. Beside the potential health effect, the results presented here are suggesting that the transmission of bacteria from breast milk and the nutrients supplied by it at this critical time point in infant’s development, could affect the colonization window of specific bacterial genera, and depending on delivery mode and breastfeeding duration, this may lead to disturbances in the oral microbial succession patterns that persist over time.

**The effect of antibiotics intake on microbiota development**

The clinical data of this cohort allowed us to assess the influence of antibiotics intake in early life (first and second year) on developing microbiota. The antibiotics courses given were mainly due to early otitis media (in 89% of cases) and included Amoxicillin (34 % of cases) and Phenoxymethylpenicillin (42 % of cases) (Table S1). Upon comparing the microbial succession in children who did or did not take antibiotics during the first two years of life, significantly divergent colonisations were observed at 24 months and 7 years of age, whereas bacterial composition at earlier time points were overlapping in children treated with antibiotics (Fig. 4a). Multivariate analyses were also performed, considering the effect of time and different confounding factors on microbiota composition. Antibiotic use had a significant effect on microbiota composition once the effects of caries status and time were removed (p=0.05) and a significant antibiotic by time interaction was found (p=0.008). There was a lower effect of antibiotics on microbiota composition (p=0.067) once breastfeeding length was included in the analysis, suggesting that part of the significance is due to the strong effect of breastfeeding on microbiota composition.

By comparing the most dominant genera (>1% of total microbiota) present in these two groups, the genus *Granulicatella* was higher in abundance at 24 months of age (p=0.003) in children not taking antibiotics while *Prevotella* (p=0.020) was more prevalent at 7 years of age in children treated with antibiotics in early life. The data suggest that the abundance of commensal genera such as *Granulicatella* (Aas et al., 2005) may be disturbed by antibiotics use while the presence of other genera, like *Prevotella*, which has been associated with several oral diseases (Aas et al., 2008), may be favoured.

In order to obtain deeper insight of microbiota alterations upon antibiotics intake, the microbial composition was assessed at species-level OTUs (Fig. 4b). The analysis revealed a high number of
bacteria uniquely present in children that were treated with antibiotics more than once during the first two years of life including several Actinomyces species at 2 and 7 years of age. Moreover, the presence of species belonging to Fusobacterium, Veillonella and Lactobacillus was also associated with antibiotics intake during the first two years of life in our cohort. The fact that Veillonella spp use organic acids as their only carbon source strongly suggests that the oral microbiota of those children is more acidogenic. On the contrary, Neisseria and Streptococcus mitis/dentisani, were present in our samples at significantly higher levels in 7-year old children that did not take antibiotics. Thus, although a divergent microbiota does not necessarily imply a negative effect for health, most significant changes in microbial composition detected in the current study as a consequence of antibiotic administration, have previously been associated with oral diseases (Alcaraz et al., 2012; López-López et al., 2017; Nyvad et al., 2013; Kolenbrander et al., 2006; Yasukawa et al., 2010; Badet & Thebaud, 2008; Bradshaw & Marsh, 1998) and future studies will need to specifically address whether antibiotic use during infancy has an effect on oral health.

It is of course possible that the divergent microbial succession patterns observed at 7 years of age might be affected by further antibiotics courses and other influencing factors, occurring during the remaining five years. However, given that the first years of age appear to represent a crucial period of microbiota development and immune system modulation and that early changes in ecological successions are those with the largest impact on community development (Amarasekare & Possingham, 2001), it is important to consider that early antibiotic treatment can have long-term consequences for microbiota development. It has to be emphasized that in adults, the original salivary microbial composition appears to be restored after antibiotic use (Zaura et al., 2015), suggesting resilience of the oral microbiome; in children, long-term alterations of the gut microbiome, as a consequence of early antibiotic administration, have been proposed to have negative effects for systemic health, including obesity and allergy (Ajslev et al., 2011; Reynolds & Finlay, 2017). Thus, the impact of early intake of antibiotics for human health deserves consideration.

**Oral microbiota in health and disease**

Caries development did not appear to be related with bacterial diversity (Fig. 1c) or bacterial load (Fig. SI1c) during the first 7 years of life. Although there were no differences between children staying caries-free and children developing caries at age 9, the density of bacteria was increasing significantly with age, probably reflecting the influence of environmental interactions and the emergence of teeth. The overall species richness was higher in children that remained caries-free by
9 years of age, but the difference was not significant (Fig. 1c). However, the potential association of lower bacterial diversity to caries risk should be further studied, as a lower bacterial diversity has been associated to caries in cross sectional studies (Simón-Soro et al., 2013). A factor reducing the possible association of caries status to diversity could be the use of saliva samples, which provide a good representation of overall oral microbial diversity but may not fully correlate with bacterial composition at the tooth biofilm, where the disease takes place (Mira, 2017).

Caries development at 9 years of age was preceded by divergent bacterial composition at 24 months of age, reaching the maximum at 7 years (Fig. 2c). At early age, no differences between caries-experienced and caries-free children were detected, suggesting that the colonization patterns and ecological factors favouring caries development are associated with later age. A critical period may include the eruption of primary incisors, primary molars and permanent first molars, where cariogenic bacteria like Mutans streptococci can adhere through glucan binding proteins (Law et al., 2007). Although these caries-linked species are considered associated to hard-tissues, there are studies suggesting that they can be acquired at any time from under 6 months (prior to first tooth eruption) to over 3 years of age (Wan et al., 2001a, 2001b). Taken together, the data here suggest that different bacterial colonization patterns were present between caries-free children and children that developed caries, however they were significant only after the second year of age.

**Bacterial composition and caries development**

Since no significant differences were observed between caries-free and caries-active children at the genus taxonomic level (Fig. SI4) and given that the genus *Streptococcus* was highly abundant in the infants’ oral cavity, it was of great interest to investigate if there were any specific Streptococci species associated with caries development in the cohort. The genus *Streptococcus* comprises a large number of species that can have positive effects on human health and some of them have started to be used as probiotics in oral diseases (Gruner et al., 2016). The OTUs found corresponded to *S. mitis/infantis/dentisani* (identical in the sequenced 16S rRNA region), *S. salivarius*, *S. sanguinis*, *S. lactarius*, *S. cristatus* and *S. mutans* (Fig. 5). *S. mitis/infantis/dentisani* were the most prevalent OTUs (ranging from 75-85%) and no difference was found between the children who did or did not develop caries at 9 years of age. *S. infantis* belongs to the *Streptococcus mitis* group (Zbinden et al., 2015) and has been associated with oral health as it significantly decreases during caries progression in the young permanent dentition (Gross et al., 2010). *S. dentisani* is a bacterial species associated with good oral health and it has been isolated from caries-free individuals (López-López et al., 2017). Because of the high sequence similarity within the *Streptococcus* genus PCR-amplified region used for Illumina sequencing, we could not distinguish which 16s rRNA
reads belonged to *S. mitis*, *S. infantis* or *S. dentisani*. To clarify this, qPCR amplification with *S. dentisani*-specific primers was performed in order to determine the acquisition of this species through age. The quantities of *S. dentisani* were undetectable by qPCR during the first year of age, suggesting that the colonization of this species might be dependent of teeth eruption. This is in agreement with its normal association with dental plaque (López-López et al., 2017). The levels of *S. dentisani* were higher in children remaining caries-free at 9 years of age in comparison with caries-experienced children, but the difference was not significant (Fig. S15).

*Streptococcus salivarius* was another commonly found species in children’s saliva (Fig. 5). Its abundance was highest at 3 months of age, ranging between 10-15% of the total streptococcal species, and decreasing steadily through time, likely opposing teeth eruption. This pioneer colonizer and a prominent member of the oral microbiota of the healthy mouth has been detected hours after birth because of its unique ability to adhere and colonize tongue and cheek mucosa (Nelson-Filho et al., 2013). Although *S. salivarius* has been intended for use as a probiotic targeting the oral cavity (Burton et al., 2006), no differences in abundance levels of this species through age were discovered between children who did or did not develop caries at 9 years of age, perhaps due to its absence from dental plaque (López-López et al., 2017). *Streptococcus lactarius* was another species encountered in infant’s saliva, predominantly at 3 and 6 months of age, to later decrease and even disappear. This species was isolated from breast milk of healthy mothers (Martín et al., 2011), explaining its high abundance in early age when the majority of the children in this cohort were breastfed. Given the long-term impact of breastfeeding for microbiota development (Fig. 2b), it is plausible that early colonization with *S. lactarius*, acquired from mother’s breast milk, could benefit later colonization by other beneficial microbial species. However, the potential role of *S. lactarius* in health and disease has not been evaluated to date.

Colonization of *Streptococcus sanguinis* started between 6 and 12 months of age and followed a similar pattern of development between children who did and did not develop caries. This species is believed to play a benign role in the oral cavity and it has been described to colonize in association to tooth emergence, at a median age of 9 months (Caufield et al., 2000). Moreover, *S. sanguinis* is recognized for its antagonistic role in dental caries since it may compete with cariogenic mutans streptococci for colonization sites on tooth surfaces (Caufield et al., 2000). Interestingly, although at very low levels, the cariogenic *S. mutans* was detected in the oral cavity of the infants already at an early age, possibly acquired through their mothers as shown before, (Law et al., 2007) with a trend of significantly higher levels at 7 years (p=0.06) in children developing caries. This is in line with previous studies where proportions of *S. mutans* in saliva were higher in children with caries when
compared to those who stayed caries-free (Lif Holgerson et al., 2015). Thus, although this species is considered mainly an inhabitant of hard tissues, our data show that it can be detected before tooth eruption and therefore the oral health of mothers and caretakers during infancy may play an important role in the transmission of this pathogen. However, S. mutans has also been detected in caries-free populations and not in all cases of childhood caries, suggesting that other species may be cariogenic pathogens (Aas et al., 2008; Law et al., 2007). In this study, children developing caries had significantly higher abundance of Streptococcus cristatus already at 3 months (p=0.026) and 24 months of age (p=0.033), compared to the children that stayed caries-free until 9 years of age. Given that S. cristatus, among other species, has been associated with severe early childhood caries (Tanner et al., 2011), even in the absence of Streptococcus mutans, its role as an important cariogenic species and potential caries risk biomarker should be further studied. Nevertheless, it must be emphasized that streptococci are extremely similar in their 16S rRNA gene sequence, particularly at the V3-V4 region analysed in the current work, and therefore the suggested association of S. cristatus with caries development should be confirmed by species-specific methodology with higher discriminatory power, like qPCR with specific primers and probes (Coffey & Shlossman, 2016).

If the association between S. cristatus and dental caries is confirmed, it must be born in mind that this species has been found to interrupt the formation of P. gingivalis biofilms by repressing the production of several virulence factors in this major periodontal pathogen (Ho et al., 2017). In our dataset, a scatterplot of the relative frequencies of Porphyromonas and S. cristatus shows an L-shape (correlation p-value for the hyperbolic regression was p=0.057), suggesting potential antagonistic behaviour (Fig. SI6), a feature that has been demonstrated in subgingival plaque samples from adults (Wang et al., 2009). Given that most Porphyromonas sequences in our samples corresponded to P. catoniae (P. gingivalis accounted only for 13.4% of total Porphyromonas reads by 7 years of age), the potential antagonism between S. cristatus and P. gingivalis may not be apparent until a later age.

In addition, LefSe analyses were performed in order to examine potential biomarkers for early caries diagnosis. No specific group of species/genera at early age could be associated with caries development at 9 years of age (data not shown), suggesting that other ecological determinants including host interactions with microbiota, play a crucial role and should be integrated in caries risk assessment models (Mira et al., 2017; Young & Featherstone, 2010). Interestingly, even though the supplementation with L. reuteri during the first year of life has been associated with reduced caries prevalence at 9 years of age (Stensson et al., 2014), no differences in caries development
related to this Lactobacilli could be detected in the present study. Given that some of the infants included in the study developed allergies during their early childhood (see Table 1), the groups were balanced according to allergy status and no relationship was found between allergies and caries onset. Even though mode of delivery and breastfeeding until 12 months of age have been shown to impact oral microbiota development in this study, no correlation between delivery mode or breastfeeding duration with dental caries could be detected. However, this could be due to low statistical power of the groups compared. Although microbiota composition clearly differed at 7 years of age between caries-free and caries-experienced children (Fig. 2c), the absence of robust individual biomarkers of caries risk at an earlier age underlines that microbial-based early diagnostic tests should not be based on single species, and new potential bacterial risk indicators should be identified (E. Kanasi et al., 2010), including S. cristatus as proposed above. Given the enormous inter- and intra-individual variability in bacterial composition at caries lesions (Simón-Soro et al., 2015), and the multi-factorial nature of oral diseases where microbial, environmental and host-associated variables are involved, a holistic, ecological approach to caries risk assessment where information about the host, the habits (including the diet and oral hygiene) and the microbes are integrated will likely provide a better estimate of caries prediction (Belda-Ferre et al., 2015; Mira et al., 2017; Young & Featherstone, 2010).

CONCLUSIONS

Only limited information is available on oral microbiome development in infants, and most studies have focused on taxonomic analysis. Thus, functional, metagenomic analyses are pending to fully understand the microbial contribution to oral health and disease (Mira, 2018). Previous studies addressing oral microbiota development in early life have been hampered by retrospective approaches, small sample sizes, lack of deep sequence coverage, limited period of follow-up and analyses at single time points. The current study demonstrates that the infant’s oral cavity gets colonized by microorganisms in a timely manner, increasing in complexity with time. In general, the data presented in the current manuscript is consistent with a model where microbiota development follows an ecological succession (Van Best et al., 2015).

In this scenario, several early colonizing species pave the way for the settlement of other microorganisms, which further expand microbial diversity towards a mature community which is more robust and resilient to change, partly because of the developed immune tolerance (Zaura et al., 2014). The presence of several species (particularly S. cristatus) at an early age was associated in this study to a higher frequency of dental caries at 9 years of age. Therefore, these findings open the possibility to use this species, together with others identified in other studies, as potential
biomarkers of caries risk. The oral cavity is a complex and heterogeneous ecosystem with many variables influencing microbial composition and function. Several external factors appear to strongly influence microbiota development, including mode of delivery, which had a short-term effect, and others like breastfeeding length or antibiotic treatment, which appeared to have a long-term impact. It is interesting to note that, on the contrary, the oral microbiome composition in adults appears to be extremely resilient to antibiotic treatment (Zaura et al., 2015). This highlights that developmental milestones that are critical for oral microbiota succession occur in particular during infancy, and that an appropriate microbial colonization pattern can be instrumental for future health. Thus, microbial exposure, feeding habits and medical interventions during those initial and fragile stages may have a lifelong impact on general microbiome composition, and their potential consequences for human health should be carefully studied.

**Acknowledgements**

We would like to acknowledge the technical assistance performed by Ann-Marie Fornander and Camilla Janefjord. We would also like to thank Alba Boix Amorós and Sandra Garcia Esteban for their great assistance in the laboratory work.

Supplementary information is available at ISME’s website.

**REFERENCES**


FIGURE LEGENDS

Fig. 1. Species richness and diversity of total microbiota in infant saliva samples. Bacterial richness and diversity (here presented by Chao1 and Shannon indices), obtained at different time points from birth to 7 years of age, were determined by 16s rRNA Illumina sequencing and OTU clustering at 97% sequence identity. Panel a) shows species richness and diversity in infants delivered vaginally (VD) or by caesarean section (C-section) at 3 months (N_{VD}=62; N_{CS}=11), 6 months (N_{VD}=72; N_{CS}=11), 12 months (N_{VD}=59; N_{CS}=10), 24 months (N_{VD}=56; N_{CS}=10) and 7 years of age (N_{VD}=68; N_{CS}=12). Panel b) represents species diversity and richness in infants breastfed for 12 months (BF) and in infants breastfed for less than 6 months of age (nBF). Analysed samples were collected at 12 months (N_{nBF}=52; N_{BF}=17), 24 months (N_{nBF}=50; N_{BF}=16), and 7 years of age (N_{nBF}=59; N_{BF}=21). Panel c) shows species richness and diversity in children developing caries (CA) and children staying caries-free (CF) during the first 9 years of life. Saliva samples were collected at 3 months (N_{CF}=40; N_{CA}=26), 6 months (N_{CF}=43; N_{CA}=31), 12 months (N_{CF}=37; N_{CA}=26), 24 months (N_{CF}=35; N_{CA}=24) and 7 years of age (N_{CF}=45; N_{CA}=30). Data are presented with means and standard errors. (*p <0.05; Mann-Whitney U-test).

Fig. 2. Salivary microbiota patterns through children’s development. Constrained correspondence analysis (CCA), here used to emphasize variations in microbiota species-level profiles, show the relationship between groups in total microbiota composition at different time points. The percentage of variation explained by constrained correspondence components is indicated on the axes. a) Microbial pattern differences in saliva from infants delivered vaginally (VD) or by caesarean section (CS), p=0.0016, at 3 months (N_{VD}=62; N_{CS}=11), 6 months (N_{VD}=72; N_{CS}=11), 12 months (N_{VD}=59; N_{CS}=10), 24 months (N_{VD}=56; N_{CS}=10) and 7 years of age (N_{VD}=68; N_{CS}=12). b) is showing microbial composition pattern differences in infants who were breastfed for 12 months (BF) and in infants breastfed (nBF) for less than 6 months (p=0.0017). The numbers of children were: 12 months (N_{nBF}=52; N_{BF}=17), 24 months (N_{nBF}=50; N_{BF}=16) and 7 years of age (N_{nBF}=59; N_{BF}=21). c) Microbial composition patterns in children developing caries (CA) and children staying caries-free (CF) during the first 9 years of life (p=0.0018). Saliva samples were collected at 3 months (N_{CF}=40; N_{CA}=26), 6 months (N_{CF}=43; N_{CA}=31), 12 months (N_{CF}=37; N_{CA}=26), 24 months (N_{CF}=35; N_{CA}=24) and 7 years of age (N_{CF}=45; N_{CA}=30). Numbers accompanying the variables (delivery mode, breastfeeding and caries onset) are representing the time points plotted. p values for CCA plots were determined by Adonis analysis (a nonparametric statistical method, R package Vegan) and significant values indicate that the factor provided can explain part of the total variability.
**Fig. 3.** Microbiota composition of dominant bacterial genera in children being or not being breastfed until 12 months of age. a) Genera classified as early colonizers. b) Genera considered to “constantly increase” are already present at 3-6 months of age at >1% frequency, and are increasing in relative proportion with time. c) Bacterial genera considered as “Late colonizers”, are defined as those present at 3-6 months of age below 1% relative abundance, which undergo a significant increase after 12 months of age. Plots are showing the relative abundance of dominant bacterial genera, as determined by Illumina sequencing of 16S rRNA gene, in saliva samples collected at 3 months, 6 months, 12 months (N\text{BF}=52; N\text{BF}=17), 24 months (N\text{BF}=50; N\text{BF}=16) and 7 years (N\text{BF}=59; N\text{BF}=21) of age. Asterisks indicate cases where p <0.05 by both Mann-Whitney U-test and Wilcoxon analysis.

**Fig. 4.** Accumulative antibiotics effect on salivary microbiota development. Panel a) shows salivary microbiota patterns in children treated with antibiotics during the first two years of life and children not taking any antibiotics during the first 2 years of life; p =0.017. p values for CCA plots were determined by Adonis analysis (a nonparametric statistical method, R package Vegan) and indicate that the factor provided can explain part of the total variability. Panel b) represents the influence of antibiotics intake on bacterial species distribution at 24 months and 7 years of age, here represented with a Venn’s diagram. The numbers indicate unique species found in children not treated with antibiotics (purple) and children treated with antibiotics (grey) and OTUs differentially distributed are stated in the tables. OTUs presented were filtered according to sequence length (>300 bp) and identity (>97% nucleotide similarity). Analysed sample sizes were: 3 months (N\text{NO}=28; N\text{YES}=11), 6 months (N\text{NO}=33; N\text{YES}=13), 12 months (N\text{NO}=29; N\text{YES}=11), 24 months (N\text{NO}=26; N\text{YES}=10) and 7 years (N\text{NO}=35; N\text{YES}=11); NO= no antibiotics intake for the first 2 years of age; YES= antibiotic consumption during the first 2 years of age.

**Fig. 5.** Relative abundance of most prevalent *Streptococci* species found in saliva samples of children staying caries-free and children developing caries during the first 9 years of age. Plots represent average relative abundance of Streptococci through time. Taxonomy assignments were performed with RDP classifier at 100% nucleotide identity. All data are presented as means with standard errors. Asterisks indicate cases where p <0.05 by both Mann-Whitney U-test and Wilcoxon analysis.
Accumulative antibiotics effect

CCA1 (22.05%)
CCA2 (15.48%)

Acinetobacter baumannii
Actinomyces oris
Actinomyces sp.
Actinobaculum sp.
Bergeyella sp.
Cardiobacterium hominis
Corynebacterium diphtheriae
Dolosigranulum pigrum
Gemella morbillorum
Gemella sanguinis
Haemophilus parainfluenzae
Haemophilus parahaemolyticus/sputorum
Kingella oralis
Lactobacillus gasseri
Leptotrichia buccalis
Neisseria elongata
Pseudomonas fluorescens
Veillonella atypica
Veillonella parvula
Selenomonas sp.
Veillonella parvula
Selenomonas sp.
Veillonella parvula
Streptococcusagalactiae/sanguinis
Streptococcus parasanguinis
Streptococcus oralis
Streptococcus sp.

Time point
3M
6M
12M
24M
7Y

Antibiotics
No antibiotics

Parvimonas sp.
Selenomonas artemidis
Stomatobaculum sp.
Aggregatibacter aphrophilus
Haemophilus haemolyticus
Neisseria oralis

4 80 26
24 months

4 76 17
7 years
**SUPPLEMENTARY INFORMATION**

**a)**

![Graph](image1)

**Log (bacterial cells/ml saliva)**

3M, 6M, 12M, 24M, 7Y

Cesarean section
Vaginally delivered

**b)**

![Graph](image2)

**Log (bacterial cells/ml saliva)**

3M, 6M, 12M, 24M, 7Y

Partially breastfed
Non breastfed

**c)**

![Graph](image3)

**Log (bacterial cells/ml saliva)**

3M, 6M, 12M, 24M, 7Y

Caries free
Caries active

**Fig. S11.** Bacterial load of the total microbiota in infant saliva samples. Bacterial density using salivary samples that were obtained at different time points until 7 years of age, was determined by qPCR using universal primers targeting 16s rDNA bacterial gene. Figures are showing bacterial density depending on the mode of delivery a), partial breastfeeding until 12 months of age b) and dental caries status c). Data are presented with median with interquartile ranges. Delivery mode: 3 months (N\(_{VD}=62\); N\(_{CS}=11\)), 6 months (N\(_{VD}=72\); N\(_{CS}=11\)), 12 months (N\(_{VD}=59\); N\(_{CS}=10\)), 24 months (N\(_{VD}=56\); N\(_{CS}=10\)) and 7 years of age (N\(_{VD}=68\); N\(_{CS}=12\)); Partial breastfeeding at 12 months: 12 months (N\(_{abf}=52\); N\(_{bf}=17\)), 24 months (N\(_{abf}=50\); N\(_{bf}=16\)) and 7 years of age (N\(_{abf}=59\); N\(_{bf}=21\)); Caries status: 3 months (N\(_{CS}=40\); N\(_{CA}=26\)), 6 months (N\(_{CF}=43\); N\(_{CA}=31\)), 12 months (N\(_{CF}=37\); N\(_{CA}=26\)), 24 months (N\(_{CF}=35\); N\(_{CA}=24\)) and 7 years of age (N\(_{CF}=45\); N\(_{CA}=30\)). M=months; Y=years. (*p <0.05, **p<0.01; Mann-Whitney U-test).
Fig. SI2. Oral microbiota development through time. Plots show the relative abundance of dominant bacterial genera at 3, 6, 12, 24 months and 7 years of age.
Fig. SI3. Microbiota composition of the most dominant bacterial genera in children delivered vaginally or through caesarean section. a) Genera considered as early colonizers. b) Genera considered to “Constantly increase” are already present at 3-6 months of age, at >1%, and they are increasing with time. c) Bacterial genera considered as “Late colonizers”, present at 3-6 months of age but below 1% of abundance and significant increase after 12 months of age. Plots are showing the relative abundance of dominant bacterial genera, as determined by Illumina sequencing of 16S rDNA gene, in saliva samples collected at 3 months (N_{VD}=62; N_{CS}=11), 6 months (N_{VD}=72; N_{CS}=11), 12 months (N_{VD}=59; N_{CS}=10), 24 months (N_{VD}=56; N_{CS}=10) and 7 years of age (N_{VD}=68; N_{CS}=12); M=months; Y=years. (*p <0.05; Mann-Whitney U-test).
**Fig. S14.** Microbiota composition of the most dominant bacterial genera in children developing dental caries and children staying caries-free until 9 years of age. a) Genera considered as early colonizers. b) Genera that are constantly increasing through the age. Plots are showing the relative abundance of dominant bacterial genera, as determined by Illumina sequencing of 16S rDNA gene, in saliva samples collected at 3 months (N_{CF}=40; N_{CA}=26), 6 months (N_{CF}=43; N_{CA}=31), 12 months (N_{CF}=37; N_{CA}=26), 24 months (N_{CF}=35; N_{CA}=24) and 7 years of age (N_{CF}=45; N_{CA}=30). CF: caries free; CA: caries active; M=months; Y=years. (*p <0.05; Mann-Whitney U-test).

**Fig. S15.** *Streptococcus dentisani* levels in infant saliva samples determined by qPCR quantification. Detection limit was established >2*10^2 cells/ml saliva. Data are presented with median and interquartile ranges. 24 months (24M): N_{CF}=15; N_{CA}=6; 7 years (7Y): N_{CF}=26; N_{CA}=19.
Fig. SI6. Correlation between *Streptococcus cristatus* and *Porphyromonas* spp in children’s salivary samples. Scatterplot shows the relative proportions of the two bacteria, which approximate a hyperbolic regression (p=0.057), suggesting potential antagonism between the two taxa. Samples with 0% proportion were removed from the analysis.

Table SI1. Accumulative antibiotics intake during the first two years of life.

<table>
<thead>
<tr>
<th>Patient code</th>
<th>Antibiotics courses (age in months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Phenoxymethylpenicillin (9), Phenoxymethylpenicillin (10), Amoxicillin (11), Phenoxymethylpenicillin (22)</td>
</tr>
<tr>
<td>2</td>
<td>Phenoxymethylpenicillin (7), Phenoxymethylpenicillin (18)</td>
</tr>
<tr>
<td>3</td>
<td>Erythromycin (9), Phenoxymethylpenicillin (13), Loracarbef (17)</td>
</tr>
<tr>
<td>4</td>
<td>Phenoxymethylpenicillin (9), Ampicillin (10), Phenoxymethylpenicillin (20), Amoxicillin (21)</td>
</tr>
<tr>
<td>5</td>
<td>Phenoxymethylpenicillin (3), Amoxicillin (3), Amoxicillin (10), Amoxicillin (13), Flucloxacillin (15), Amoxicillin (15), Amoxicillin (19), Amoxicillin (24)</td>
</tr>
<tr>
<td>6</td>
<td>Phenoxymethylpenicillin (10), Phenoxymethylpenicillin (12), Phenoxymethylpenicillin (23)</td>
</tr>
<tr>
<td>7</td>
<td>Phenoxymethylpenicillin (21), Benzyl-penicillin and Phenoxymethylpenicillin (23)</td>
</tr>
<tr>
<td>8</td>
<td>Phenoxymethylpenicillin (8), Amoxicillin (9), Phenoxymethylpenicillin (16)</td>
</tr>
<tr>
<td>9</td>
<td>Phenoxymethylpenicillin (6), Amoxicillin (7), Amoxicillin (10), Amoxicillin (11), Phenoxymethylpenicillin (19), Amoxicillin (21), Amoxicillin (22)</td>
</tr>
<tr>
<td>10</td>
<td>Bactrim (6), Furadantin (6-11, as prophylaxis), Phenoxymethylpenicillin (20)</td>
</tr>
<tr>
<td>11</td>
<td>Phenoxymethylpenicillin (2), Amoxicillin (3), Phenoxymethylpenicillin (13), Amoxicillin (13)</td>
</tr>
<tr>
<td>12</td>
<td>Ampicillin (11), Erythromycin (11), Amoxicillin (14), Erythromycin (17), Erythromycin (19)</td>
</tr>
</tbody>
</table>