

# Loss of microbial topography between oral and nasopharyngeal microbiota and development of respiratory infections early in life



# Loss of microbial topography between oral and nasopharyngeal microbiota and development of respiratory infections early in life

Wing Ho Man[1,2†], Melanie Clerc[3†], Wouter A.A. de Steenhuijsen Piters[1,3,4§], Marlies A. van Houten[2], Mei Ling J.N. Chu[1,4], Jolanda Kool[5], Bart J.F. Keijser[5,6], Elisabeth A.M. Sanders[1], Debby Bogaert[1,3*]

**Affiliations:**

[1] Department of Paediatric Immunology and Infectious Diseases, Wilhelmina Children's Hospital/University Medical Center Utrecht, Utrecht, The Netherlands;

[2] Spaarne Gasthuis Academy, Hoofddorp and Haarlem, The Netherlands;

[3] Medical Research Council/University of Edinburgh Centre for Inflammation Research, Queen's Medical Research Institute, University of Edinburgh, Edinburgh, United Kingdom;

[4] Department of Medical Microbiology, University Medical Center Utrecht, Utrecht, The Netherlands;

[5] Microbiology and Systems Biology Group, Netherlands Organisation for Applied Scientific Research (TNO), Zeist, The Netherlands;

[6] Department of Preventive Dentistry, Academic Center for Dentistry Amsterdam (ACTA), University of Amsterdam, Amsterdam, The Netherlands.

† These authors contributed equally to this work.

§ Present address: Department of Pulmonary Medicine, St. Antonius Hospital, Nieuwegein, The Netherlands.

* To whom correspondence should be addressed. E-mail: D.Bogaert@ed.ac.uk

Wing Ho Man, MD

Department of Paediatric Immunology and Infectious Diseases

Wilhelmina Children's Hospital/University Medical Center Utrecht

P.O. Box 85090, 3508 AB Utrecht, The Netherlands

winghoman@gmail.com

Tel: +31 23 2241665

Melanie Clerc, PhD

Medical Research Council/University of Edinburgh Centre for Inflammation Research

Queen's Medical Research Institute, University of Edinburgh

47 Little France Crescent, EH16 4TJ, Edinburgh, United Kingdom

M.Clerc@ed.ac.uk

Tel: +44 131 2426582

Wouter A.A. de Steenhuijsen Piters, MD

Medical Research Council/University of Edinburgh Centre for Inflammation Research

Queen's Medical Research Institute, University of Edinburgh

47 Little France Crescent, EH16 4TJ, Edinburgh, United Kingdom

Wouter.deSteenhuijsen@ed.ac.uk

Tel: +44 131 2426582

Marlies A. van Houten, MD, PhD

Spaarne Gasthuis Academy

P.O. Box 900, 2000 VB Haarlem, The Netherlands

MvanHouten2@spaarnegasthuis.nl

Tel: +31 23 2241665

Mei Ling J.N. Chu, BSc

Department of Paediatric Immunology and Infectious Diseases

Wilhelmina Children's Hospital/University Medical Center Utrecht

P.O. Box 85090, 3508 AB Utrecht, The Netherlands

M.L.J.N.Chu@umcutrecht.nl

Tel: +31 88 7557635

Jolanda Kool, BSc

Microbiology and Systems Biology Group

Netherlands Organisation for Applied Scientific Research (TNO)

Utrechtseweg 48, 3704 HE Zeist, The Netherlands

Jolanda.kool@tno.nl

Tel: +31 88 8661407

Bart J.F. Keijser, PhD

Microbiology and Systems Biology Group

Netherlands Organisation for Applied Scientific Research (TNO)

Utrechtseweg 48, 3704 HE Zeist, The Netherlands

bart.keijser@tno.nl

Tel: +31 88 8665141

Elisabeth A.M. Sanders, MD, PhD

Department of Paediatric Immunology and Infectious Diseases

Wilhelmina Children's Hospital/University Medical Center Utrecht

P.O. Box 85090, 3508 AB Utrecht, The Netherlands

L.Sanders@umcutrecht.nl

Tel: +31 30 2742185

Debby Bogaert, MD, PhD (corresponding author)

Medical Research Council/University of Edinburgh Centre for Inflammation Research

Queen's Medical Research Institute, University of Edinburgh

47 Little France Crescent, EH16 4TJ, Edinburgh, United Kingdom

D.Bogaert@ed.ac.uk

Tel: +44 131 2426582

## Author contributions

MAvH, EAMS, and DB designed the experiments and wrote the study protocols. MAvH was responsible for clinical data collection. MLC was responsible for sample preparation, and MLC and BK for 16S-rRNA gene amplicon sequencing. WHM, MC, WAAdSP and DB were responsible for bioinformatic processing and statistical analyses, and wrote the paper. All authors significantly contributed to interpreting the results, critically revised the manuscript for important intellectual content, and approved the final manuscript.

## Funding

This work was supported in part by The Netherlands Organisation for Health Research and Development (ZonMW; grant 91209010); The Netherlands Organisation for Scientific research (NWO-VIDI; grant 91715359); Wilhelmina Children's Hospital and Spaarne Gasthuis Hoofddorp intramural funds; and Topconsortia for knowledge and innovation (Agri & Food; TKI-AF-12190).

## Running Title

Loss of topography precedes respiratory infections.

## Impact in Two Sentences

Early-life loss of microbial topography accompanied by influx of oral taxa in the nasopharynx precedes the development of respiratory tract infections. This may lead to new insights for prevention of respiratory tract infections and antibiotic utilization in childhood.

## Descriptor

10.11 Pediatrics: Respiratory Infections

## Word count

Abstract: 259 / 250 words.

Main text: 3,755 / 3,500 words.

## At a glance commentary

*What is the current scientific knowledge on this subject?*

The microbial community composition of the nasopharynx, including the absence of gram-positive commensals, and the low-abundant presence of oral bacterial species, is strongly associated with (susceptibility to) respiratory tract infections (RTIs). It is, however, unknown whether the oral microbiota itself and/or its temporal dynamics relative to the nasopharyngeal microbiota are associated with development of RTIs

*What does this study add to the field?*

In a prospective, longitudinal birth cohort study, we characterized the oral and nasopharyngeal microbiota over the first six months of life in 112 infants both during health (nine sampling moments) and at the moment of RTIs (n = 1,750 samples). Our results clearly demonstrate that an apparent loss of microbial topography

can be observed prior to and during RTIs when paired samples are being analyzed. This loss of topography was driven by the absence of beneficial microbes, the presence and abundance of potential pathogenic bacteria, and a proportional influx of oral species in the nasopharyngeal niche on the individual's level. We unveiled bacterial biomarkers associated with loss of topography and the subsequent development of RTIs, and could also link their colonization characteristics to a known risk factor for development of RTIs, i.e. start of daycare, suggesting the microbiota represent the biological link between risk factors for and actual development of infections.

This article has an online data supplement, which is accessible from this issue's table of content online at www.atsjournals.org.

## ABSTRACT

**Rationale:** The respiratory microbiota is increasingly being appreciated as an important mediator in the susceptibility to childhood respiratory tract infections (RTIs). Pathogens are presumed to originate from the nasopharyngeal ecosystem.

**Objectives:** To investigate the association between early-life respiratory microbiota and development of childhood RTIs.

**Methods:** In a prospective birth cohort (Microbiome Utrecht Infant Study: MUIS), we characterized the oral microbiota longitudinally from birth until six months of age of 112 infants (9 regular samples/subject) and compared them with nasopharyngeal microbiota using 16S-rRNA-based sequencing. We also characterized oral and nasopharynx samples during RTI episodes in the first half year of life.

**Measurements and Main Results:** Oral microbiota were driven mostly by feeding type, followed by age, mode of delivery and season of sampling. In contrast to our previously published associations between nasopharyngeal microbiota development and susceptibility to RTIs, oral microbiota development was not directly associated with susceptibility to RTI development. However, we did observe an influx of oral taxa, such as *Neisseria lactamica*, *Streptococcus*, *Prevotella nanceiensis*, *Fusobacterium* and *Janthinobacterium lividum*, in the nasopharyngeal microbiota prior to and during RTIs, which was accompanied by reduced presence and abundance of *Corynebacterium*, *Dolosigranulum* and *Moraxella* spp. Moreover, this phenomenon was accompanied by reduced niche differentiation indicating loss of ecological topography preceding confirmed RTIs. This loss of ecological topography was further augmented by start of daycare, and linked to consecutive development of symptomatic infections.

**Conclusions:** Together, our results link the loss of topography to subsequent development of RTI episodes. This may lead to new insights for prevention of RTIs and antibiotic utilization in childhood.


**Key words**: respiratory microbiota, child, respiratory tract infections, development, risk factors.

## INTRODUCTION

Acute respiratory tract infections (RTIs) are one of the most common health problems in young children, and a major cause of morbidity, are a major reason for antibiotic prescriptions and health-care costs during childhood (1–3). Risk factors are amongst others mode of delivery, lack of breastfeeding, indoor air pollution and situations of crowding (4): however, the variation in RTI susceptibility between primarily healthy children is still largely unexplained. In this context, the composition of the microbial community in the upper respiratory tract is increasingly being appreciated as an important gatekeeper to respiratory health. The respiratory microbiota is assumed to provide colonization resistance against pathogenic microorganisms and to shape the maturing immune system in early life (5).

The respiratory tract is composed of multiple distinct ecological niches and the microbiota in the nasopharynx in particular is thought to play a key role in mediating susceptibility to RTIs (6–9). Previously, we showed a link between early life nasopharyngeal microbiota composition and subsequent susceptibility to RTIs in a prospective birth cohort study, making use of small sampling intervals (Microbiome Utrecht Infant Study [MUIS]) (6). We found that the nasopharyngeal microbiota development was accelerated in the first month of life, which was accompanied by prolonged reduction of *Corynebacterium* spp. and *Dolosigranulum* spp., decreased microbial community stability, and subsequently a higher number of RTI in the first year of life. Also, interestingly, the presence of low-abundant oral bacteria, such as *Neisseria* and *Prevotella* spp., was associated with a higher number of RTIs in the first year of life. Also, other studies showed evidence that the presence of oral species in the nasopharynx is associated with susceptibility to RTIs (6, 10, 11).

The oral microbial community may interact with nasopharyngeal microbiota and act as a microbial source for the nasopharyngeal niche. For example, *Streptococcus pneumoniae*, which is assumed to have the nasopharynx as its primary ecological niche, is equally often detected in saliva (12). The degree of interaction between oral and nasopharyngeal microbiota is however not clear as no studies so far have simultaneously investigated the concordance between nasopharyngeal and oral microbiota, and their

separate relationship with RTIs. We hypothesize that insight in the maturation and differentiation of the respiratory microbiota in infants, i.e. oral and nasopharyngeal microbiota and its relation with development of RTIs, may increase our understanding of pathogenesis of respiratory infections.

Here, we study the composition, development and topographical differentiation of paired oral and nasopharyngeal samples of 112 children of the MUIS-study cohort from birth until six months of age, in relation to development of parentally reported RTI infections in the first year of life.

## METHODS

Details on the study design, sample and data collection and bioinformatics/statistical methods can be found in the supplemental Methods. Sequence data that support the findings of this study have been deposited in the NCBI Sequence Read Archive database under accession number SRP141299.

### Data collection

The specifics on study design and inclusion criteria can be found elsewhere (13). Sequence data of nasopharyngeal samples were used previously in a study focussing on the development of the microbiota in the nasopharynx, and its relationship with early life respiratory health (6). For the current analyses, we added sequencing data obtained from the same children (n=112) representing the development of the oral microbiota in the first 6 months of life (total samples, oral cavity, n=846; nasopharynx, n=853; details of sampling in supplemental Methods). Over the course of the first 6 months of life, each child was sampled within 2 hours after birth, and on days 1, 7 and 14, followed by sampling at 1, 2, 3, 4 and 6 months of age. Further, additional samples were taken within 48 hours in case of a parent-reported symptomatic RTI, defined as presence of fever ≥38°C for >6h combined with malaise and presence of RTI symptoms. The number of RTIs ranged from 0 to 7, but due to lack of power at either end of this spectrum, we stratified the population into three groups: 0-2 RTIs, 3-4 RTIs and 5-7 RTIs, based on the distribution of RTIs in our population (6).

### 16S rRNA sequencing

DNA extraction and library preparation for the V4 region of the bacterial 16S rRNA gene was performed as previously described and detailed in the supplemental Methods (6). To avoid OTUs with identical annotations, we refer to OTUs using their taxonomical annotations combined with a rank number based on the abundance of each given OTU. The raw OTU-counts table was used for the CSS normalization required for the analysis with the *metagenomeSeq* package. The OTU-proportions table was used for all other downstream analyses.

### Statistical analysis

Benjamini-Hochberg adjusted p-values (q-values) were generated where appropriate. A p-/q-value of 0.05 was considered significant.

In order to assess differences in overall community compositions between samples at different time points, we performed non-metric multidimensional scaling (NMDS) using Bray-Curtis dissimilarities between samples (nMDS(); *vegan* package). We visualised the microbial succession patterns separately for each niche and tested for the significance of both niche (2-level factor) and time point (10-level factor) using adonis2() (a function based on permANOVA, *vegan* R-package). A similar *adonis2*-analysis was performed to assess the independent effect of lifestyle and environmental factors (i.e. breastfeeding, chronological age, season of sampling and birth mode, antibiotic use in the previous four weeks, presence of siblings, daycare attendance and in-house smoke exposure) on the overall oral microbiota composition over time.

We also investigated the stability of the temporal succession patterns at the individual level, and whether certain bacterial taxa (biomarkers) were enriched at specific time points. We used an unsupervised hierarchical clustering approach based on the Bray-Curtis dissimilarity between samples as previously described (6).

In addition, to investigate the effects of birth mode and breastfeeding on the temporal dynamics of other oral bacterial taxa, we used smoothing spline analysis of variance (SS-ANOVA, fitTimeSeries();

*metagenomeSeq* package). All models were adjusted for breastfeeding/birth mode, seasonality, siblings and daycare attendance.

To determine whether individual bacterial taxa were indicative for the oral microbiota or nasopharyngeal microbiota, we performed indicator species analysis (multipatt(); *indicspecies* package) using a strict cut-off for the indicator value (stat>0.5) (14).

To study the temporal dynamics of microbial topography, we first defined topography as the total microbiota community variance explained by niche (oral vs. nasopharynx samples) in relation to the variance explained by subject (adonis2(); *vegan* package). To estimate the robustness of our findings, we performed this analysis on 100 rarefactions at a sequencing depth of 3,000 reads. As a second measure of topography we used Bray-Curtis niche dissimilarity of paired oral and nasopharynx samples, because we also wanted to assess the actual dissimilarity between niches. Moreover, using this second method we could model the temporal dynamics of topography against exact age at sampling as a continuous variable.

## RESULTS

### Characterization of the study population

**Baseline characteristics of the study population stratified by number of RTIs experienced in the first year of life have been published previously and showed a positive association between the number of RTIs and the presence of siblings under the age of 5, daycare attendance, and the number of antibiotic courses (6).**

### Characterization of the oral and nasopharyngeal microbiota

We analyzed a total of 66,986,053 reads (median, 28,400 reads/sample; minimum, 3,134 reads/sample), with 918 operational taxonomic units (OTUs) belonging to 19 bacterial phyla which were retained after filtering. Thirty-six saliva samples were excluded because they yielded less than 3,000 reads. Especially directly after birth, the bacterial composition of oral samples and the nasopharynx showed significant

overlap indicating similar source communities. From one week of life onwards, however, the developmental trajectories of the two rapidly discern, indicating rapid niche differentiation. In oral samples, the local community remained strongly dominated by Firmicutes throughout the first 6 months of life, with *Streptococcus* (OTU rank number 1) accounting for 51% of the bacteria and no major changes in oral microbiota over time (**Figure 1A**). In the nasopharynx, Firmicutes represented by *Streptococcus* [1], *Staphylococcus* [3] and *Dolosigranulum pigrum* [5] dominated the local community during the first month of life, with a gradual increase in Actinobacteria such as *Corynebacterium propinquum* [4] from day 1 on, and Proteobacteria such as *Moraxella* [2 and 8] and *Haemophilus* [6] from week 1 on (**Figure 1B**), as previously described (6). The respective niches became over time predominantly populated by niche-specific communities, with specific oral indicator taxa (i.e. taxa indicative for a niche based on their specificity and the fidelity) like *Streptococcus*, *Veillonella*, *Rothia* and *Gemella* spp., and nasopharyngeal indicator taxa like *Corynebacterium*, *Dolosigranulum* and *Moraxella* spp. (*indicspecies* analysis, stat>0.5, Benjamini-Hochberg corrected q<0.001, **Figure E1**; full list of indicator taxa, **Table E1**).

### Oral and nasopharyngeal microbiota development over time

We used non-metric multidimensional scaling (NMDS) based on the Bray-Curtis dissimilarity index to visualize the development of the respective microbial communities over time on group level (**Figure 2**), proving microbiota development becomes more niche-dependent with time (*adonis*, niche-by-time point interaction, p=0.001).

To explore the temporal dynamics of oral and nasopharyngeal microbiota at the individual level, we applied hierarchical clustering. We found six oral microbiota clusters and six clusters for the nasopharyngeal microbiota. The succession patterns and composition of the individuals' microbiota profile over time is visualized using an alluvial diagram shown in **Figures E2A-C**. The oral microbiota was overwhelmingly dominated by a very stable *Streptococcus* [1] cluster, with a temporary increase of the *Streptococcus salivarius* [7] cluster between week two and month one. The remaining clusters were much smaller but

induced most variation over time. This in contrast to nasopharyngeal microbiota profiles, following as described previously, a more gradual developmental trajectory (6, 10). The higher temporal stability of oral microbiota clusters over nasopharyngeal microbiota was confirmed by calculating Bray-Curtis dissimilarities of all individuals' microbiota over time ($p<0.001$, **Figure E2C**).

### Environmental drivers and their impact on overall oral microbiota composition

We next assessed the association between lifestyle and environmental factors, and oral microbiota composition over time. Breastfeeding showed the strongest independent association with overall oral microbiota composition in all samples (multivariable *adonis2*, $R^2=2.8\%$, $p<0.001$), followed by age ($R^2=2.7\%$, $p<0.001$), season of sampling (0.9%, $p<0.001$) and birth mode (0.3%, $p=0.046$). Antibiotic use in the previous month, presence of siblings, day care attendance, pacifier use and second-hand smoke exposure, which are all known additional drivers of the nasopharyngeal microbiota, were not associated with oral microbiota composition ($p>0.10$), suggesting resilience of oral communities for these influences. The impact of feeding type, as well as the impact of birth mode and season on oral community composition varied over time (*adonis2*, factor-by-time point interaction, $p<0.001$) with season of birth being the most important driver directly after birth (*adonis2*, $R^2=14.2\%$, $p<0.001$), both season ($R^2 = 5.7\%$, $p=0.028$) and birth mode ($R^2=2.7\%$, $p=0.031$) at day one, and feeding type (being breastfed or not) after one week of life ($R^2=3.4\text{-}7.8\%$, $p\leq0.006$).

### Breastfeeding and birth mode influence individual oral microbial taxa

The abundance of the most abundant bacterial taxon *Streptococcus* [1] in the oral cavity but not in the nasopharynx was associated with breastfeeding (GAMM niche-by-feeding type interaction: $t=-6.4$, $p<0.001$; **Figure 3**). With respect to other taxa, we found that in oral samples *Bifidobacterium* and *Streptococcus salivarius* [7] abundances were associated with both vaginal birth and breastfeeding, whereas *Alloprevotella* abundance was associated with birth by caesarean section and formula feeding (**Figure 4**). Other oral taxa

such as different *Veillonella*, low abundant *Prevotella* and *Megasphaera* spp. were associated with vaginal birth, and *Neisseria lactamica* was associated with birth by caesarean section only (multivariable *fitTimeSeries*, q<0.10). Breastfeeding was additionally associated with early and prolonged enrichment (beyond month 3) of *Staphylococcus*, *Haemophilus* and *Lactobacillus* spp. in oral microbiota (q<0.02), and formula feeding was associated with enrichment of *Prevotella melaninogenica*, *Prevotella nanceiensis*, *Granulicatella*, *Leptotrichia*, *Fusobacterium*, *Rothia* spp. and other streptococci (multivariable *fitTimeSeries*, q<0.01; **Figures E3 and E4**, **Tables E2 and E3**).

### Oral microbiota maturation in relation to susceptibility to RTI

In a previous report on the early life development of the nasopharyngeal microbiota, we observed that microbial maturation was significantly associated with the number of observed RTIs over the first year of life (6). For oral microbiota, we observed no association between the number of RTIs and microbiota age, nor with α-diversity (all associations p>0.05, **Figure E5**). In contrast to the nasopharyngeal microbiota (6), the oral microbiota composition at one month of life was also not predictive of the number of RTIs over the first year of life (*adonis*, p=0.69). These results suggest that the nasopharyngeal microbiota remains the key community driving respiratory health and infections.

### Topography of the upper respiratory microbial communities in relation to RTI susceptibility

We further studied whether the topographical differentiation between the two niches over time, defined by the level of topographical differentiation as the total microbiota community variance explained by niche (oral vs. nasopharynx samples) versus the variance explained by subject (*adonis2*-analysis assessing the marginal effects of niche and subject for each timepoint) over time. Directly after birth and at day one, topographical differentiation between the two niches had not yet occurred and the majority of the variance was explained by subject rather than by niche ($R^2$>75% and ~5%, respectively; **Figure 5A**). Thereafter, a clear gradual increase in topographical differentiation was observed over the first three months of life (niche

differentiation at age three months, $R^2$=38.4%, p<0.001), while variance explained by subject became less important and was no longer significant after the first week of life (p>0.05). However, during a RTI, the two niches became more similar again with a marked decline in variance explained by niche at time of infection ($R^2$=20.3%, p<0.001) and a significant increase in variance explained by subject again ($R^2$=56.6%, p=0.002; **Figure 5B**). This seems to indicate that during a RTI episode, there is loss of topography between the two niches. Interestingly, however, the loss of topography was observed already well before the confirmed RTI episode occurred, and the variance explained by subject became already significant at the first time-point preceding the confirmed RTI, in general one month before the RTI episode occurred (average -28 days; $T_{RTI}$=-1; subject, $R^2$=37.7%, p=0.010; niche, $R^2$ 37.6%, p<0.001; **Figure 5B**), suggesting loss of topography might contribute to dysbiosis and development of symptomatic infections.

We found similar results when applying Bray-Curtis dissimilarity between paired oral and nasopharyngeal samples as alternative measure of topographical differentiation. Again, after the first week of life, the dissimilarity between niches increased gradually until the age of three months. When the birth cohort was stratified into groups of children with up to two RTI episodes (healthy reference group) versus children with more than two RTI episodes over the first year of life, we found that the healthy reference group had a significantly higher niche dissimilarity from 8 weeks on compared to children who eventually suffered from more than two RTIs (p=0.032; **Figure 5C**). Also, during RTI episodes, the niche dissimilarity had decreased when compared to the reference group (p=0.016). This loss of topography during RTI episodes became more pronounced over time (**Figure 5C**).

The difference between the healthy reference group and the group with more than two RTI episodes was already observed in early infancy and well before the first infections occurred, with most RTIs occurring after the age of 4 months. This suggests that the loss of topography precedes RTI episodes rather than coincides or follows RTIs. Loss of topography seems therefore indicative of factual development of an RTI. This was confirmed by studying niche dissimilarity directly before and after a confirmed RTI in individual children; we again found that loss of topography preceded the confirmed RTI, with a significantly reduced

329 niche dissimilarity approximately one month before the RTI episode but not yet two months before the
330 confirmed RTI ($T_{RTI}$=-1 vs. $T_{RTI}$=-2, p=0.04; **Figure 5D**).

331 Five children developed a RTI within one month of start of daycare. We observed that these five
332 children had a significantly higher loss of topography following start of daycare compared to children who
333 did not develop a RTI within one month after start of daycare ($T_{daycare}$=Start daycare, p=0.016), suggesting
334 again that loss of topography reflects microbial instability, and thereby increases the risk of subsequent
335 development of RTI following start of daycare. However, loss of topography was apparent already one
336 month before start of daycare ($T_{daycare}$=-2 vs. $T_{daycare}$=-1, p=0.032, **Figure 6A**), indicating that daycare
337 attendance may have enhanced the loss of topography that preceded the development of a RTI episode, but
338 that this is interdependent with the level of topographical differentiation in respiratory microbiota to start
339 with. Overall, these results provide evidence for the hypothesis that loss of topography is a proxy for
340 respiratory microbiota at disequilibrium, and that loss of topography facilitates symptomatic RTI
341 development after encountering a second trigger, such as viral or bacterial pathobiont exposure at daycare
342 (15).

343 **Loss of topography is driven by enrichment of oral taxa in the nasopharynx**

344 When looking in more detail at loss of topography in all children and associated bacterial taxa that lead to
345 less differentiation between the niches, the loss of topography appeared to be primarily driven by an influx
346 of oral indicator taxa (predominantly *Streptococcus* [1]) in nasopharyngeal samples ($T_{RTI}$=-2 vs. $T_{RTI}$=-1,
347 median combined relative abundance 1.4% vs. 7.3%, p=0.03; **Figure E6AB**), but not the reversed, i.e. the
348 increase of nasopharyngeal indicator taxa in oral samples was not observed (**Figure E6C**). Similarly, we
349 observed an increase of oral indicator taxa in nasopharyngeal samples in all children well before the start of
350 daycare ($T_{daycare}$=-2 vs. $T_{daycare}$=-1, median combined relative abundance 0.8% vs. 2.3%, p=0.012), that
351 remained high at the start of daycare ($T_{daycare}$=-1 vs. $T_{daycare}$=Start daycare, 2.3% vs. 6.8%, p=0.43). This
352 influx of oral taxa was more pronounced in the five children who developed a RTI within one month after

start of daycare as compared to those who did not report RTI symptoms within the first month (**Figure 6B**). On individual bacterial taxon level, these five most susceptible children had significantly higher nasopharyngeal abundances of the oral indicator taxa *Neisseria lactamica*, *Streptococcus* [18], *Prevotella nanceiensis*, *Fusobacterium* and *Janthinobacterium lividum* already before start of daycare (**Figure E7**). Noticeably, these oral indicator taxa were not only enriched in the nasopharynx of the five children mentioned above but were in general present in higher abundance in the oral microbiota of children who were not breastfed for at least 3 months of life or were born by caesarian section (**Tables E1 and E2**). In contrast, the children, who appeared protected against developing a RTI following start of daycare attendance, showed higher nasopharyngeal abundances of several nasopharyngeal indicator taxa like *Corynebacterium*, *Dolosigranulum* and *Moraxella* spp. before the start of daycare (**Figure E7**). Together, the study data support the hypothesis that the nasopharyngeal microbiota acts as primary gatekeeper to maintain respiratory health. The early presence and abundance of beneficial biomarker taxa in the nasopharynx that are associated with breastfeeding and vaginal delivery, increase stability, and prevent loss of topography and subsequent RTIs. In contrast, early life loss of topography in the upper respiratory tract accompanied by influx in the nasopharynx of oral taxa, which is associated with caesarean section and formula feeding, may instigate (susceptibility to) RTIs.

## DISCUSSION

It is becoming increasingly apparent that human respiratory health is significantly mediated by the microbial communities that reside in the upper respiratory tract. We and others have recently described the temporal dynamics of the nasopharyngeal microbiota and its relationship with (susceptibility to) RTIs (6–9). Studies investigating early life oral microbiota development are, however, sparse (16, 17). Moreover, although the presence of oral taxa in the nasopharynx is associated with RTI presence and severity (6, 9, 11), longitudinal studies linking nasopharyngeal and oral microbiota dynamics, and their relationship with RTIs have not yet been performed.

Here, we provide substantial new knowledge to the current evidence base. We demonstrate that the infant oral microbiota is almost immediately dominated by *Streptococcus* spp. and is very stable throughout the first six months of life. Additionally, oral microbiota was highly associated with feeding type but showed no direct relationship with (susceptibility to) RTIs. However, loss of upper respiratory microbial topography, driven by a proportional influx of oral taxa in the nasopharyngeal niche, appears to precede RTI episodes. Moreover, daycare attendance, which is a well-known risk factor for development of RTI episodes due to high exposure to pathobionts (4, 15), induced a further loss of topography but this depended on the level of topographical differentiation prior to the start of daycare attendance. Together, this implies that the nasopharyngeal microbiota composition and differentiation is linked to RTI susceptibility.

Our findings linking the loss of topography to subsequent development of RTI episodes corroborate a previous report on the loss of microbial topography in the upper respiratory tract of elderly adults, who are, similar to young children, more susceptible to respiratory infections than mid-aged-adults (18). In line with our results, the authors demonstrated in elderly a replacement of the nasal community as observed in mid-aged adults by an oropharyngeal-like population of microbes, which are characterized by an increase in the abundance of *Streptococcus* spp. This further signifies the importance of microbial topography and the potential role of oral bacteria in perpetuating inflammation when beyond the oral niche. In a recent *in vivo* study in mice colonization of oral bacteria in the intestine significantly induced $T_H1$ cells and led to severe intestinal inflammation (19). Interestingly, this occurred only in susceptible hosts, e.g. mice with antibiotic-induced intestinal dysbiosis, but not in healthy mice, suggesting that loss of microbial topography after a second trigger occurs more easily in ecosystems lacking stability induced by keystone species (19). Similarly, we found that the loss of upper respiratory microbial topography preceding a RTI covaried with the loss of nasopharyngeal abundance of presumed beneficial commensals such as *Corynebacterium*, *Dolosigranulum* and *Moraxella* spp. (5, 20–22), all species with high susceptibility to routine antibiotic treatment (5). Eradication of these species by inappropriate use of broad-spectrum antibiotics therefore seems undesirable (3). Also, caution is warranted with respect to new vaccination strategies that may affect

presence of *Moraxella* species (23). Efforts should instead be made to uphold these beneficial commensals to promote respiratory health. Some promising steps have already been taken; nasal application of *Corynebacterium* spp. has been proven to provide resistance against RSV and secondary pneumococcal pneumonia in infant mice (24) and appeared safe in small pilot studies involving healthy human adults (25, 26).

## ACKNOWLEDGMENTS:

We gratefully acknowledge A.A.T.M. Bosch, G. Biesbroek and the research team of the Spaarne Gasthuis Academy for their execution of and unrelenting dedication to the study; the department of Obstetrics and Gynecology of the Spaarne Gasthuis and participating midwifery clinics for recruiting participants for this study; the laboratory staff of the University Medical Center Utrecht and TNO for their lab work and the cooperating institutes for their commitment to the project. Most of all, we are indebted to all the participating children and their families.

**Competing interests:**

No conflict of interest related to the present study. EAMS declares to have received unrestricted research support from Pfizer, grant support for vaccine studies from Pfizer and GSK and fees paid to the institution for advisory boards or participation in independent data monitoring committees for Pfizer and GSK. DB declares to have received unrestricted fees paid to the institution for advisory work for Friesland Campina and well as research support from Nutricia and MedImmune. No other authors reported financial disclosures. Funding sources had no role in the study design, in the collection, analysis and interpretation of data, in writing the report, and the decision to submit the paper for publication. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

## REFERENCES


1. Liu L, Oza S, Hogan D, Chu Y, Perin J, Zhu J, Lawn JE, Cousens S, Mathers C, Black RE. Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet* 2016;388:3027–3035.
2. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015;386:743–800.
3. Donnelly JP, Baddley JW, Wang HE. Antibiotic Utilization for Acute Respiratory Tract Infections in U.S. Emergency Departments. *Antimicrob Agents Chemother* 2014;58:1451–1457.
4. Rudan I. Epidemiology and etiology of childhood pneumonia. *Bull World Health Organ* 2008;86:408–416.
5. Man WH, de Steenhuijsen Piters WAA, Bogaert D. The microbiota of the respiratory tract: gatekeeper to respiratory health. *Nat Rev Microbiol* 2017;15:259–270.
6. Bosch AATM, Piters WAA de S, van Houten MA, Chu MLJN, Biesbroek G, Kool J, Pernet P, de Groot P-KCM, Eijkemans MJC, Keijser BJF, Sanders EAM, Bogaert D. Maturation of the Infant Respiratory Microbiota, Environmental Drivers, and Health Consequences: A Prospective Cohort Study. *Am J Respir Crit Care Med* 2017;196:1582–1590.
7. Teo SM, Mok D, Pham K, Kusel M, Serralha M, Troy N, Holt BJ, Hales BJ, Walker ML, Hollams E, Bochkov YA, Grindle K, Johnston SL, Gern JE, Sly PD, Holt PG, Holt KE, Inouye M. The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development. *Cell Host Microbe* 2015;17:704–15.
8. Vissing NH, Chawes BLK, Bisgaard H. Increased risk of pneumonia and bronchiolitis after bacterial colonization of the airways as neonates. *Am J Respir Crit Care Med* 2013;188:1246–1252.
9. De Steenhuijsen Piters WAA, Heinonen S, Hasrat R, Bunsow E, Smith B, Suarez-Arrabal MC, Chaussabel D, Cohen DM, Sanders EAM, Ramilo O, Bogaert D, Mejias A. Nasopharyngeal microbiota, host transcriptome, and disease severity in children with respiratory syncytial virus infection. *Am J Respir Crit Care Med* 2016;194:1104–1115.
10. Biesbroek G, Tsivtsivadze E, Sanders EAM, Montijn R, Veenhoven RH, Keijser BJF, Bogaert D. Early Respiratory Microbiota Composition Determines Bacterial Succession Patterns and Respiratory Health in

Children. *Am J Respir Crit Care Med* 2014;190:1283–1292.

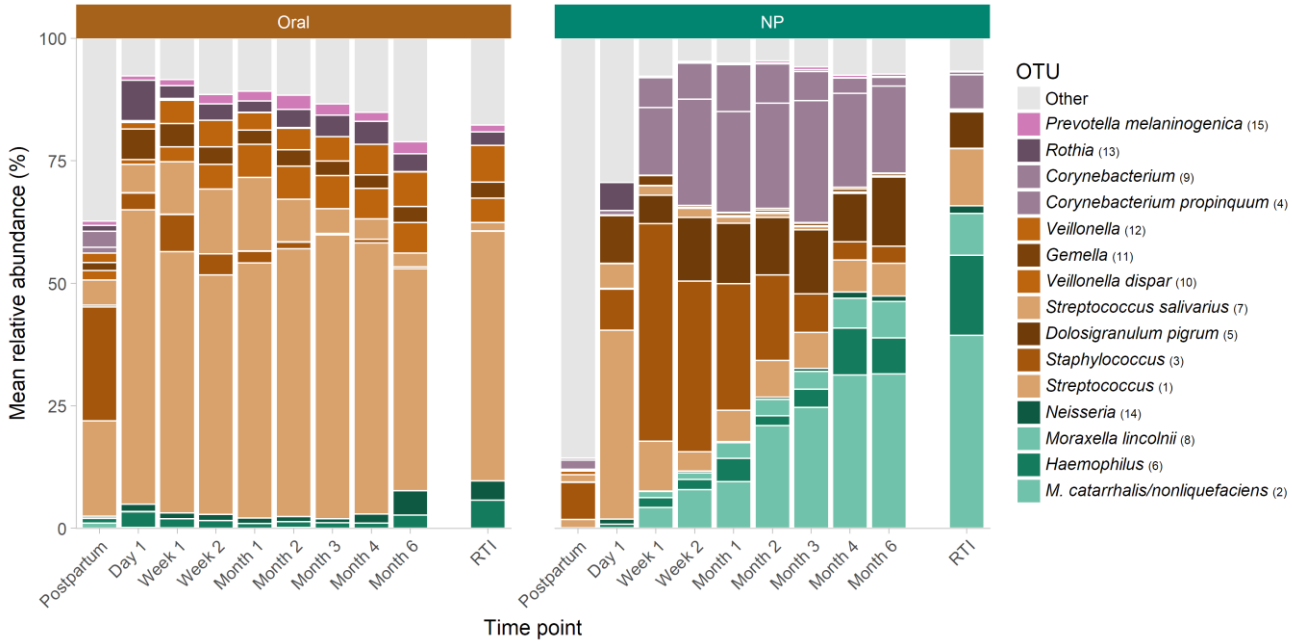
11. Lappan R, Imbrogno K, Sikazwe C, Anderson D, Mok D, Coates H, Vijayasekaran S, Bumbak P, Blyth CC, Jamieson SE, Peacock CS. A microbiome case-control study of recurrent acute otitis media identified potentially protective bacterial genera. *BMC Microbiol* 2018;18:13.
12. Wyllie AL, Chu MLJN, Schellens MHB, van Engelsdorp Gastelaars J, Jansen MD, van der Ende A, Bogaert D, Sanders EAM, Trzciński K. Streptococcus pneumoniae in saliva of Dutch primary school children. *PLoS One* 2014;9:e102045.
13. Bosch AATM, Levin E, van Houten MA, Hasrat R, Kalkman G, Biesbroek G, de Steenhuijsen Piters WAA, de Groot P-KCM, Pernet P, Keijser BJF, Sanders EAM, Bogaert D. Development of Upper Respiratory Tract Microbiota in Infancy is Affected by Mode of Delivery. *EBioMedicine* 2016;9:336–345.
14. Cáceres M De, Legendre P. Associations between species and groups of sites: indices and statistical inference. *Ecology* 2009;90:3566–3574.
15. Nesti MMM, Goldbaum M. Infectious diseases and daycare and preschool education. *J Pediatr (Rio J)* 2007;83:299–312.
16. Gomes-Filho IS, Passos JS, Da Cruz SS. Respiratory disease and the role of oral bacteria. *J Oral Microbiol* 2010;2:.
17. de Steenhuijsen Piters WAA, Sanders EAM, Bogaert D. The role of the local microbial ecosystem in respiratory health and disease. *Philos Trans R Soc B Biol Sci* 2015;370:20140294.
18. Whelan FJ, Verschoor CP, Stearns JC, Rossi L, Luinstra K, Loeb M, Smieja M, Johnstone J, Surette MG, Bowdish DME. The loss of topography in the microbial communities of the upper respiratory tract in the elderly. *Ann Am Thorac Soc* 2014;11:513–521.
19. Atarashi K, Suda W, Luo C, Kawaguchi T, Motoo I, Narushima S, Kiguchi Y, Yasuma K, Watanabe E, Tanoue T, Thaiss CA, Sato M, Toyooka K, Said HS, Yamagami H, Rice SA, Gevers D, Johnson RC, Segre JA, Chen K, Kolls JK, Elinav E, Morita H, Xavier RJ, Hattori M, Honda K. Ectopic colonization of oral bacteria in the intestine drives TH1 cell induction and inflammation. *Science* 2017;358:359–365.
20. Ramsey MM, Freire MO, Gabrilska RA, Rumbaugh KP, Lemon KP. Staphylococcus aureus Shifts toward Commensalism in Response to Corynebacterium Species. *Front Microbiol* 2016;7:1230.
21. Bomar L, Brugger SD, Yost BH, Davies SS, Lemon KP. Corynebacterium accolens Releases

Antipneumococcal Free Fatty Acids from Human Nostril and Skin Surface Triacylglycerols. *MBio* 2016;7:e01725-15.

22. Aebi C. *Moraxella catarrhalis – Pathogen or Commensal?* In: Curtis N, Finn A, Pollard AJ, editors. *Adv Exp Med Biol* New York, NY: Springer New York; 2011.
23. Murphy TF, Parameswaran GI. Moraxella catarrhalis, a Human Respiratory Tract Pathogen. *Clin Infect Dis* 2009;49:124–131.
24. Kanmani P, Clua P, Vizoso-Pinto MG, Rodriguez C, Alvarez S, Melnikov V, Takahashi H, Kitazawa H, Villena J. Respiratory Commensal Bacteria Corynebacterium pseudodiphtheriticum Improves Resistance of Infant Mice to Respiratory Syncytial Virus and Streptococcus pneumoniae Superinfection. *Front Microbiol* 2017;8:1613.
25. Uehara Y, Nakama H, Agematsu K, Uchida M, Kawakami Y, Abdul Fattah ASM, Maruchi N. Bacterial interference among nasal inhabitants: eradication of Staphylococcus aureus from nasal cavities by artificial implantation of Corynebacterium sp. *J Hosp Infect* 2000;44:127–133.
26. Kiryukhina N V., Melnikov VG, Suvorov A V., Morozova YA, Ilyin VK. Use of Corynebacterium pseudodiphtheriticum for elimination of Staphylococcus aureus from the nasal cavity in volunteers exposed to abnormal microclimate and altered gaseous environment. *Probiotics Antimicrob Proteins* 2013;5:233–238.

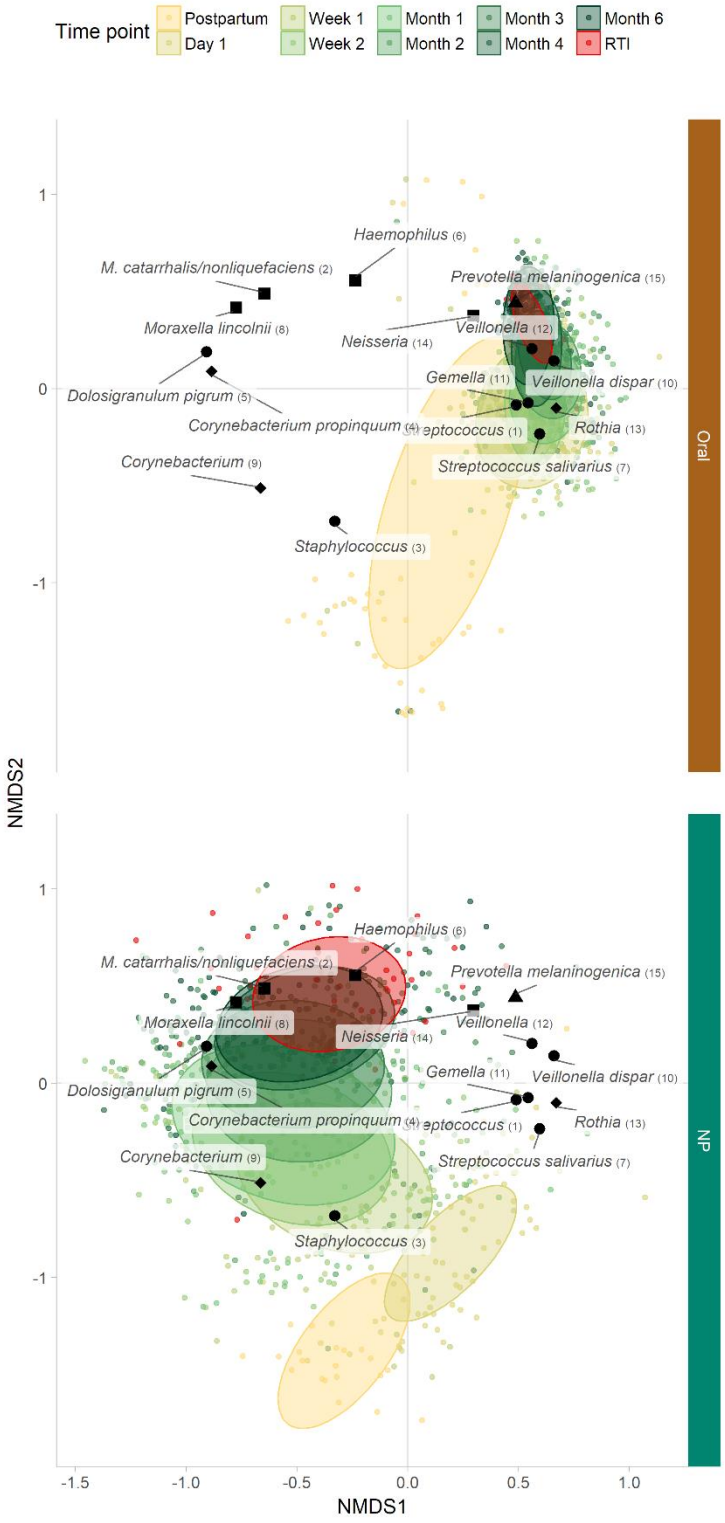
# FIGURES

**Figure 1. Mean relative abundance of the 15 most abundant shared bacterial taxa.**



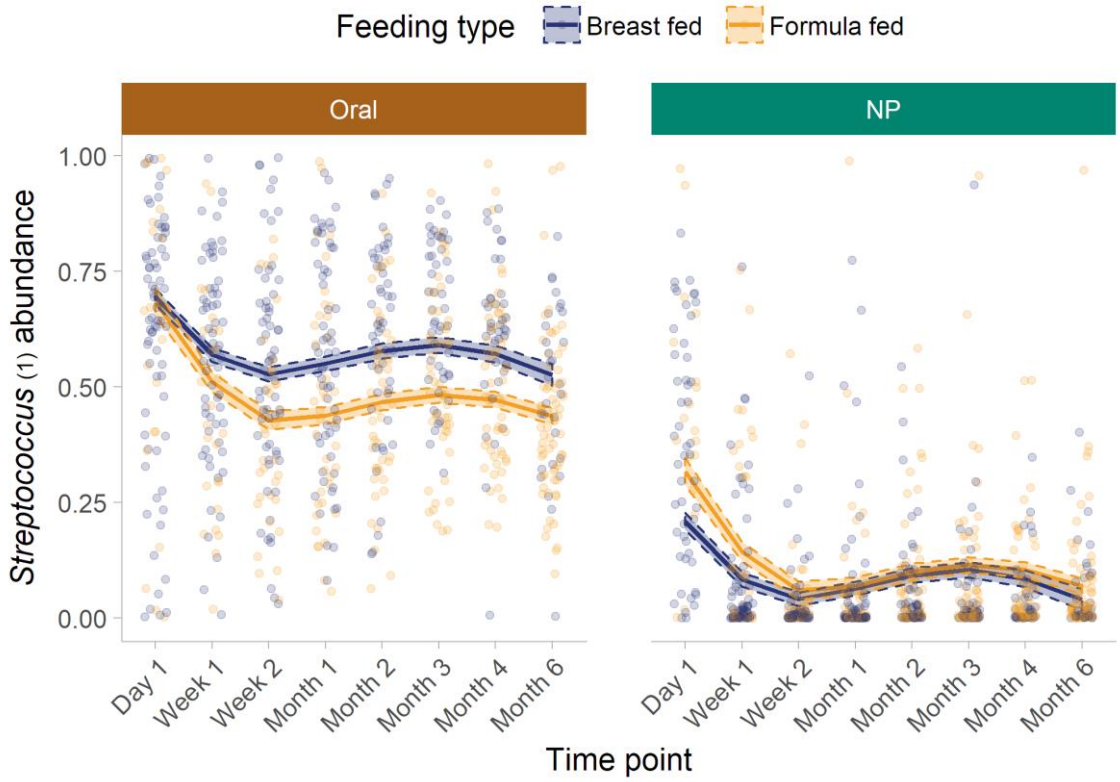

Stratified by niche: oral (left panel, n per timepoint: Postpartum 58; Day 1 89; Week 1 92; Week 2 96; Month 1 101; Month 2 98; Month 3 99; Month 4 98; Month 6 96; RTI 15) and nasopharynx (NP) samples (right panel, n per timepoint: Postpartum 21; Day 1 68; Week 1 108; Week 2 111; Month 1 112; Month 2 109; Month 3 111; Month 4 105; Month 6 108; RTI 55). The taxonomical annotation of bacterial taxa is combined with a rank number based on the abundance of each given taxon. Bacterial taxa that were not among the 15 highest ranking were collapsed and referred to as “Other”. Color shades represent phylum affiliation of individual bacterial taxa: brown = Firmicutes, green = Proteobacteria, purple = Actinobacteria, pink = Bacteroidetes, grey = residuals.

**Figure 2. NMDS plots visualizing niche-specific microbiota succession patterns during the first six months of life.**



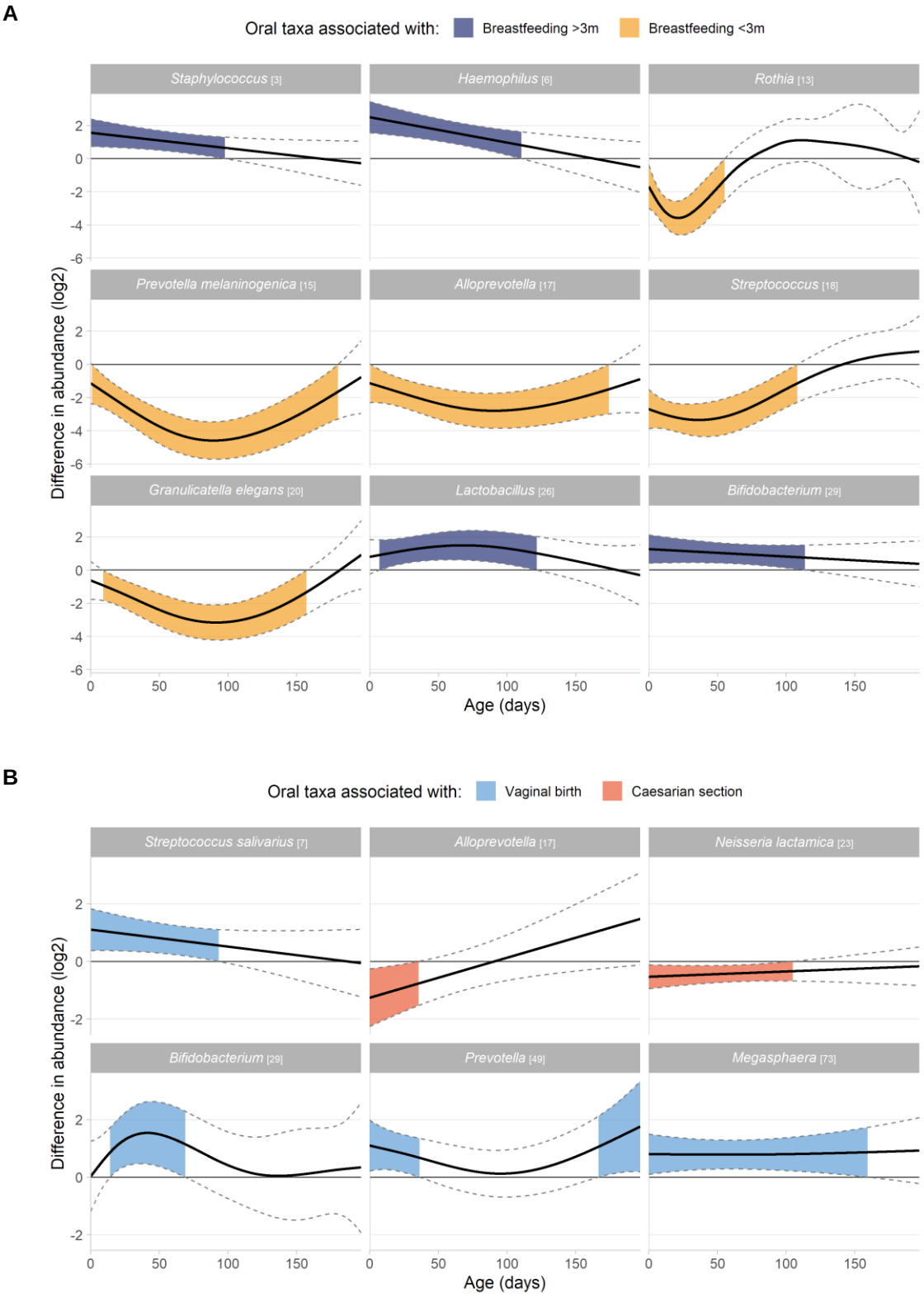
The ordination based on Bray-Curtis dissimilarity index is split by niche: oral samples (upper panel), nasopharynx (NP, lower panel). Each point represents the microbial community composition of a single sample. Samples taken during health (oral, n=829; nasopharynx, n=798) are colored based on the time point at which they were taken (yellow [day 0] to dark green [month 6]). In addition, samples taken during RTI are depicted in red (oral, n=17; nasopharynx, n=55;). Ellipses represent the standard deviation around each group of samples. Black symbols represent the 15 most abundant bacterial taxa (based on mean relative abundance across niches/time points), with the shape depicting the phylum affiliation (circle = Firmicutes, square = Proteobacteria, diamond = Actinobacteria, triangle = Bacteroidetes). The taxonomical annotation of these bacterial taxa is combined with a rank number based on the abundance of each given taxon. Stress = 0.21. RTI = respiratory tract infection. Ordination data for nasopharyngeal microbiota of this cohort were published previously *(6)*.

**Figure 3. The effect of feeding type on *Streptococcus* [1] dynamics over time.**



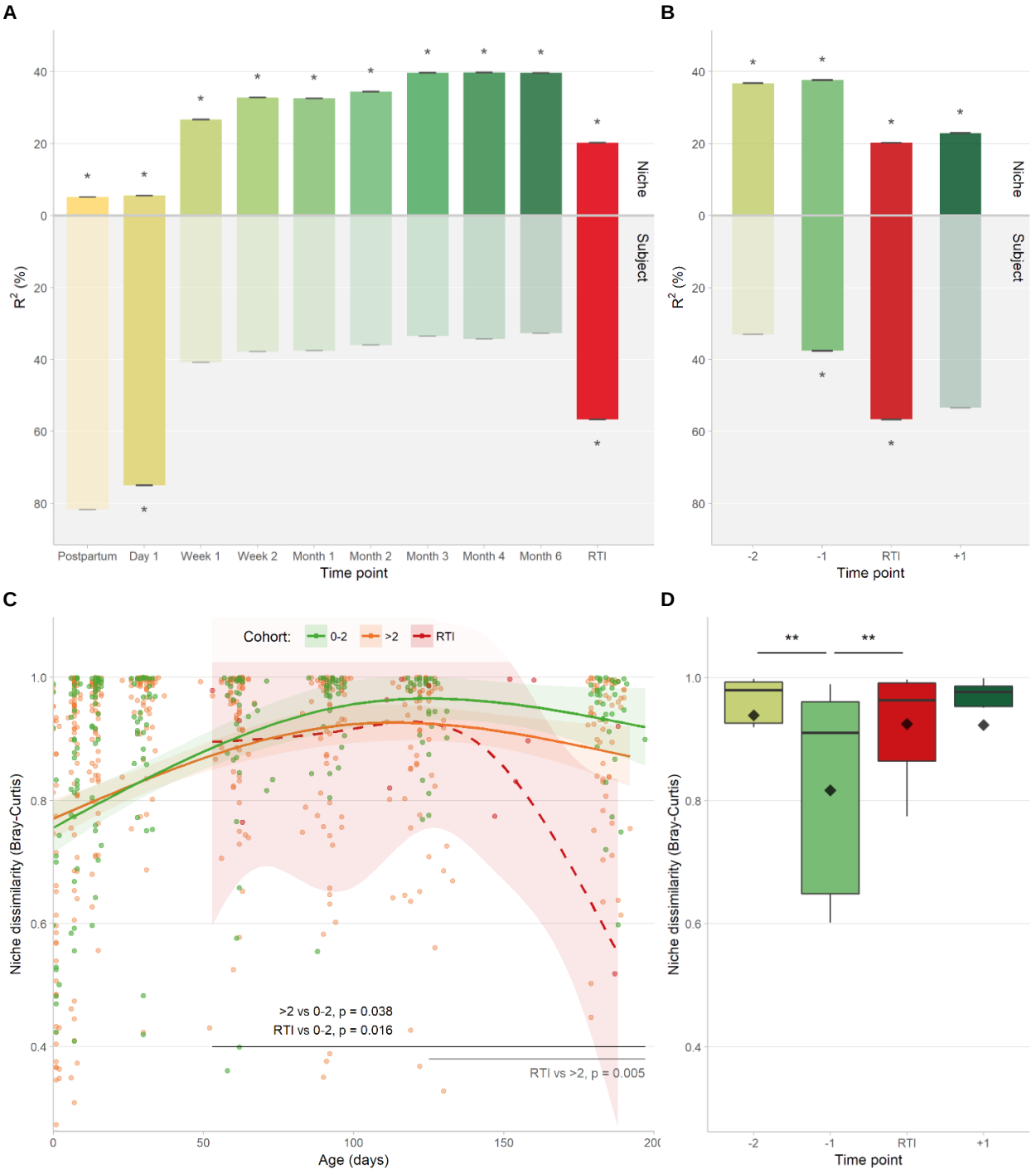

Breastfeeding is the most important driver of oral microbiota: depicted are the abundances of the main biomarker spp. of saliva, i.e. *Streptococcus* [1], stratified by mode of feeding for both the oral (left panel) and nasopharynx samples (right panel). Dots represent individual data, whereas lines represent GAMM model predictions with standard error (shaded areas). In oral samples, *Streptococcus* [1] was significantly enriched in breastfed children (blue) compared to formula-fed children (yellow). In the nasopharynx, *Streptococcus* [1] abundances were very similar between the two feeding types except during the first week of life where formula-fed children had higher *Streptococcus* [1] abundance than breastfed children.

531 **Figure 4. Temporal associations of bacterial taxa in oral samples with feeding type and birth mode.**



**(A)** $Log_2$ difference in abundance of bacterial taxa in oral samples (solid lines) and 95% confidence intervals (dashed lines) between children exclusively breast-fed until the age of 3 months and those who were not exclusively breast-fed within this same timeframe (n=48 and n=53, respectively), estimated using *fitTimeSeries*, adjusted for birth mode, season, presence of siblings, and daycare attendance. Time intervals where the bacterial taxa differ significantly between groups are colored according to whether they are increased present in infants with breastfeeding >3 months (blue) versus breastfeeding <3 months (yellow). **(B)** $Log_2$ difference in abundance of bacterial taxa in oral samples between children vaginally born and those born by caesarian section (n=62 and n=39, respectively), estimated using *fitTimeSeries*, adjusted for feeding type, season, presence of siblings, and daycare attendance. Time intervals where bacteria differ significantly between groups are colored according to whether they are increased in vaginal birth (blue) versus caesarean section (red). The taxonomical annotation of bacterial taxa is combined with a rank number based on the relative abundance of each given taxon (i.e. 1 is the most abundant taxa, followed by 2, 3, 4 etc.). Only bacterial taxa with q<0.10 are depicted. Figure **(A)** and **(B)** are a snippet of **Figure E3** and **E4**, respectively.

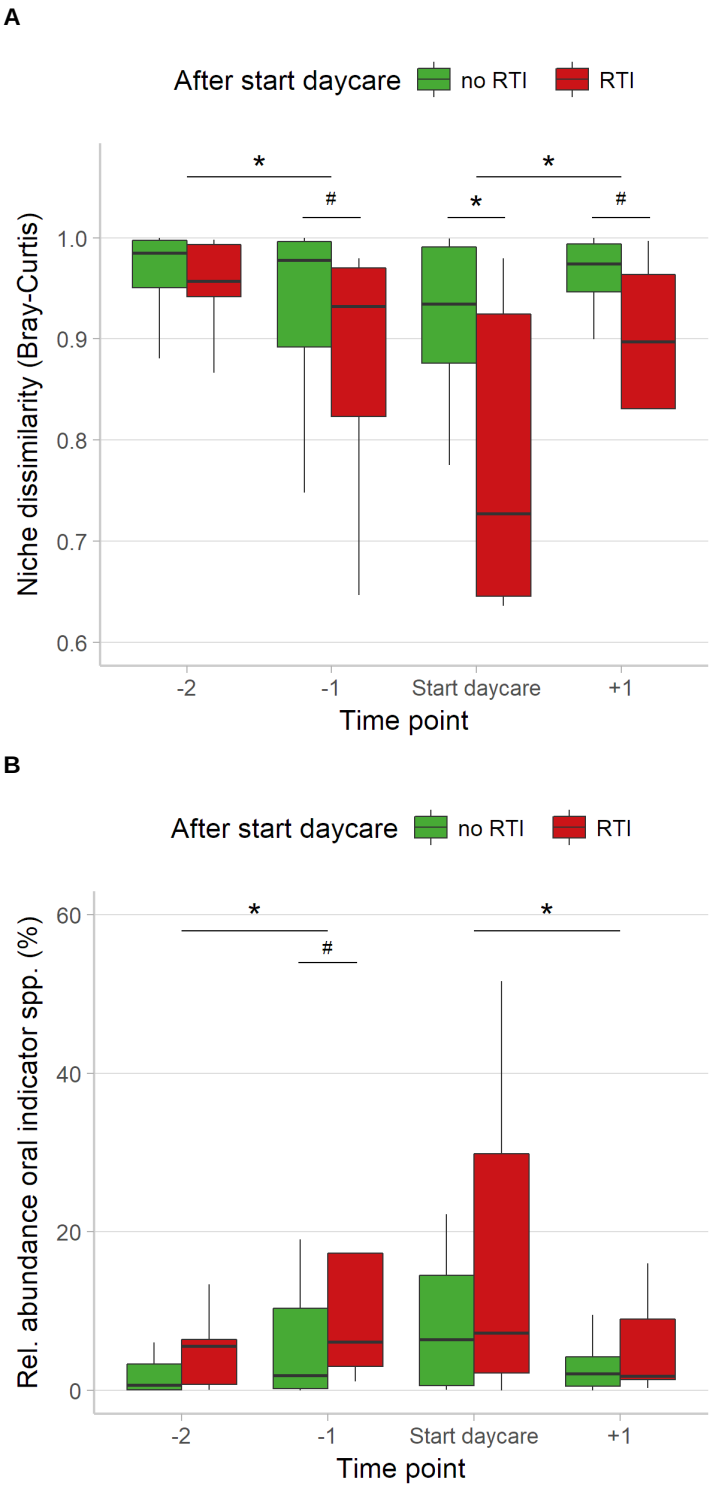
**Figure 5. Temporal dynamics of microbial topography.**




We studied the temporal dynamics of topographical differentiation in relation to RTIs using **A** and **B** *adonis2* ($R^2$; variance explained by niche relative to subject) and **C** and **D** by Bray-Curtis dissimilarity metrics. **(A)** For each sampling time point (n=981; colours ranging from yellow [day 0] to dark green [month 6], and red [RTI]) we calculated the topographical differentiation defined as the variance in total microbiota composition explained ($R^2$) by niche vs. the

549 variance in total microbiota composition explained ($R^2$) by subject (*adonis2*-analysis per time point). Bars represent the
550 average $R^2$ explained over 100 random rarefactions of the microbiota data. Error bars represent standard deviations.
551 Significant associations between either niche and/or subject and microbiota composition are depicted as solid bars and
552 are annotated with an asterisk, whereas non-significant results are depicted as transparent bars without asterisks.
553 Directly after birth and at the first day, differentiation regarding microbial topography is not yet apparent. This is reflected
554 by a big (and significant) $R^2$ for subject (bottom half) and a small $R^2$ for niche (top half). Over time the topography
555 becomes more differentiated as an increasing part of the total microbiota variation is explained by niche (i.e. the $R^2$ bar
556 for niche becomes larger over time) over that explained by subject (i.e. the $R^2$ bar for subject becomes smaller and not
557 significant). However, during a RTI the topography seems to be lost, i.e. a larger and significant part of the total
558 microbiota variation is again explained by subject.
559 **(B)** Because the children had RTIs at different ages, we also calculated differentiation regarding microbial topography
560 using the same definition as in **(A)** stratified by time point relative to a RTI (n=44 paired samples; light green shades
561 [before RTI], red [during RTI], dark green [after RTI]). '-2' represents two time points before RTI (56±8 [mean±sd] days
562 prior to RTI), '-1' represents one time point before RTI (28±7 days prior to RTI), 'RTI' represents time of RTI (age =
563 123±37 days) and '+1' represents one time point after RTI (25±16 days after RTI). This stratification of paired samples
564 in relation to time before, during and after RTI better reflects the dynamics of the topography in relation to RTI
565 development and suggests that a loss of topography already occurs before a symptomatic RTI, i.e. a larger and
566 significant part of the total microbiota variation is again explained by subject at T=-1.
567 **(C)** To confirm the above findings, we alternatively calculated the topographical differentiation between both niches in
568 a different way, i.e. using the Bray-Curtis niche dissimilarity for paired samples. We plotted this against the chronological
569 age stratified by number of RTIs (≤2 vs >2 RTIs; green and orange, respectively) experienced during the first year of
570 life. In addition, the topographical differentiation between both niches during a RTI is depicted (red). Dots represent
571 individual data points. Lines represent smooth spline fits. The shaded area around each smoothing spline represents
572 the 95% confidence interval. The confidence interval for the topographical differentiation during RTIs is very large initially
573 because of the RTIs rarely occurred before four months of age. P-values are based on a linear mixed model, including
574 age (spline) and number of RTIs as fixed effects and subject as random effect.
575 **(D)** Boxplots, including diamond shaped point indicating means, of the topographical differentiation between both niches
576 before, during and after RTI (n=44 infants, paired samples of same individuals as depicted in **[B]**, but calculating
577 topographical differentiation by Bray-Curtis dissimilarity as in **[C]**). A decline in niche dissimilarity (i.e. loss of topography)
578 was already observed (on average 28 days) prior to an RTI. P-values are based on paired Wilcoxon signed-rank tests.
579 Only '-2' vs. '-1', '-1 vs. 'RTI' and 'RTI' vs. '+1' were tested. **, $p<0.01$; *, $0.01 p<0.05$.

**Figure 6. Loss of topography coincides with enrichment of oral taxa in the nasopharynx.**




**(A)** Niche dissimilarity in relation to start daycare stratified by development of RTI. Loss of topography is already observed prior to first daycare attendance in children who rapidly (within one month) develop a RTI after start of daycare

(green, 'RTI') compared to the children who are resilient against RTIs at start of daycare (red, 'no RTI'). '-2' represents two time points before start of daycare (~63 days prior to start of daycare), '-1' represents one time point before start of daycare (~31 days prior to start of daycare), 'Start daycare' represents the first sample obtained after start of daycare (mean age = 97 days) and '+1' represents one time point after start of daycare (~35 days after start of daycare). Start of daycare appears to aggravate loss of topography rather than inducing it.

**(B)** Combined relative abundance of oral indicator taxa in nasopharyngeal samples plotted against time points relative to the start of daycare attendance (n=172). Children who did not develop an RTI shortly after start daycare had lower abundances of oral indicator taxa in their nasopharynx already prior to start of daycare compared to the ones who did develop a RTI shortly after start of daycare, suggesting daycare does not seem to be the driver of loss of topography and influx of oral taxa in the nasopharynx, rather a secondary trigger.

P-values are based on paired Wilcoxon signed-rank tests. Only '-2' vs. '-1', '-1 vs. 'Start daycare' and 'Start daycare' vs. '+1', and 'no RTI' vs. 'RTI' for each time point were tested. **, $p<0.01$; *, $0.01 p<0.05$; #, $p<0.10$.

596