EBioMedicine 46 (2019) 486–498

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

EBioMedicine
Published by THE LANCET

Research paper

Temporal association of the development of oropharyngeal microbiota with early life wheeze in a population-based birth cohort

Elizabeth A. Powell [a,1], Sara Fontanella [a,1], Eve Boakes [a], Danielle Belgrave [b], Alex G. Shaw [a], Emma Cornwell [a], Roberto Fernandez-Crespo [a], Colin G. Fink [c], Adnan Custovic [d,*], J. Simon Kroll [a]

[a] Section of Paediatrics, Department of Medicine, Imperial College London, St. Mary's Hospital Campus, London W2 1PG, UK
[b] Microsoft Research Cambridge, 21 Station Road, Cambridge CB1 2FB, UK
[c] Micropathology Ltd, The Venture Centre, Sir William Lyons Road, University of Warwick Science Park, Coventry CV4 7EZ, UK
[d] National Heart and Lung Institute, Imperial College London, St. Mary's Hospital Campus, London W2 1PG, UK

ARTICLE INFO

Article history:
Received 3 May 2019
Received in revised form 12 July 2019
Accepted 13 July 2019
Available online 25 July 2019

Keywords:
Wheeze
Microbiome
Granulicatella
Neisseria

ABSTRACT

Background: A critical window in infancy has been proposed, during which the microbiota may affect subsequent health. The longitudinal development of the oropharyngeal microbiota is under-studied and may be associated with early-life wheeze. We aimed to investigate the temporal association of the development of the oropharyngeal microbiota with early-life wheeze.

Methods: A population-based birth cohort based in London, UK was followed for 24 months. We collected oropharyngeal swabs at six time-points. Microbiota was determined using sequencing of the V3-V5 region of the 16S rRNA-encoding gene. Medical records were reviewed for the outcome of doctor diagnosed wheeze. We used a time-varying model to investigate the temporal association between the development of microbiota and doctor-diagnosed wheeze.

Findings: 159 participants completed the study to 24 months and for 98 there was complete sequencing data at all timepoints and outcome data. Of these, 26 had doctor-diagnosed wheeze. We observed significant increase in the abundance of *Neisseria* between 9 and 24 months in children who developed wheeze ($p = 0{\cdot}003$), while in those without wheezing there was a significant increment in the abundance of *Granulicatella* ($p = 0{\cdot}012$) between 9 and 12 months, and of *Prevotella* ($p = 0{\cdot}018$) after 18 months.

Interpretation: A temporal association between the respiratory commensal *Granulicatella* and also *Prevotella* with wheeze (negative), and between *Neisseria* and wheeze (positive) was identified in infants prior to one year of age. This adds to evidence for the proposed role of the microbiota in the development of wheeze.

Fund: Research funding from the Winnicott Foundation, Meningitis Now and Micropathology Ltd.

1. Introduction

A number of studies have investigated the development of the normal microbiota in the respiratory system [1] and at other sites [2], postulating that there is a critical period in early life during which the microbiota may have long-lasting effects on future health. This is supported by findings from birth cohorts [3] and animal studies [4] demonstrating that early perturbations in the microbiota increase the risk of subsequent disease, whereas perturbations later on (as early as age one year in humans) do not appear to confer comparable risk.

The presence of microbes throughout the respiratory tract, even in sites previously considered as sterile, has been established with the aid of next generation sequencing [5]. It has been proposed that microbiota of the lower airways and lungs are established by colonization from the upper airways [6]. Studies in healthy subjects have shown that lower airways are colonized (at least in part) by bacteria from the oropharynx, dominated by members of the *Firmicutes*, *Bacteroidetes*, and *Proteobacteria* phyla [7]. However, most studies to date have concentrated on the nasopharyngeal and oral microbiota, and comparatively fewer studies have investigated the role of oropharyngeal microbiota.

In a culture-based study, Bisgaard *et al* demonstrated an increased prevalence of *Streptococcus pneumoniae*, *Moraxella catarrhalis* and *Haemophilus influenzae* in hypopharyngeal secretions in one-month-old infants who subsequently developed persistent wheeze, or required hospitalization for wheeze or asthma [8]. Using a non-culture approach,

* Corresponding author.
E-mail address: a.custovic@imperial.ac.uk (A. Custovic).
[1] Joint first authors.

https://doi.org/10.1016/j.ebiom.2019.07.034

Research in context

Evidence before study

We used PubMed to interrogate the MEDLINE database, using the search terms ("Respiratory system") AND ("microbiome" OR "microbiota") AND ("wheez*" OR "asthma"), with no language or date restrictions, for papers published on or before 1 January 2019 using MESH terms where available. The search yielded 219 articles, several reports show differences in the respiratory microbiota between children with asthma compared to healthy controls. Two longitudinal studies specifically assessed the relationship between the early development of the microbiota in healthy infants and subsequent wheeze. Teo *et al,* found an association between early asymptomatic nasopharyngeal *Streptococcus* colonization and subsequent wheeze, supporting an early culture study by Bisgaard *et al*, relating asymptomatic colonization with potential respiratory pathogens including *S.pneumoniae* with subsequent wheeze. In the other, Ta *et al*, compared the early nasal microbiota in children who went on to develop rhinitis and/or wheeze to controls, finding that those infants who wheezed had an increase in *Oxalobacteraceae* and *Aerococcaceae* and a reduction in *Corynebacteriaceae* and *Staphylococcaceae*.

Added value of this study

Ours is the first study to report the longitudinal oropharyngeal microbiota in infancy and its relation to wheeze. The study adds evidence to the proposition that commensal bacteria may have a beneficial role, protecting against the development of wheeze.

Implications for all the available evidence

There is a window during infancy in which the microbiota of each niche in the upper respiratory tract matures. Strategies deployed in early infancy to modulate the microbiota or emulate its downstream effects may prevent the development of wheeze.

Teo et al. [9] have shown an increased risk of chronic wheeze by age 5 years in children who had a higher abundance of *Streptococcus* species in the nasopharyngeal microbiota at 7–9 weeks of age, but did not find the other associations reported by Bisgaard et al. [8] Recent analyses carried out in the context of a randomized placebo-controlled trial of the effect of probiotic *Lactobacillus reuteri*, have shown that children developing asthma by school age had altered oral microbiota with lower diversity of salivary bacteria and a highly divergent bacterial composition at age 7 years [10]. Furthermore, the authors observed differences in early infancy, with increased abundance of *Gemella haemolysans* in children developing allergies, and *Lactobacillus gasseri and L. crispatus* in healthy children. The longitudinal development of the nasal microbiota of infants with rhinitis and wheeze differed from those of controls [11], with an increase in abundance of *Oxalobacteraceae* and *Aerococcaceae* in children with wheeze and rhinitis. Conversely healthy infants had a higher abundance of *Staphylococcaceae* and *Corynebacteriaceae*. It has also been shown that microbiota in patients with asthma differs from that in healthy subjects (with an increase in *Moraxella* sp., and a decrease in *Prevotella* spp.) [5,12], or those with cystic fibrosis [13].

We propose that longitudinal changes in the respiratory microbiota communities in infancy and early life are related to the development of childhood wheezing illness. To test our hypothesis, we used oropharyngeal samples collected at six time points over the first two years of life to investigate the temporal association of the changes in respiratory tract microbiota with the development of early-life wheeze.

2. Methods

2.1. Study design, setting and participants

The Development Of the Respiratory Microbiota in Infants and Children (DORMICe) study was conceived to assess the development of the upper respiratory tract (URT) microbiota over the first two years of life and its relation to clinical outcomes. The study was approved by the Research Ethics Committee (12/LO/1362); parents gave written informed consent.

2.1.1. Setting

Participants were recruited at St Mary's Hospital, Imperial College NHS Trust, London (a public university hospital). Pregnant women were approached in the antenatal clinics for consideration of participation of their infant in the study from December 2012 to August 2013. Healthy babies born ≥37 weeks gestation whose parents had sufficient spoken English to maintain follow-up were included.

2.1.2. Recruitment

All healthy term (≥37 completed weeks of gestation) babies born at St Mary's Hospital between January 2013 and October 2013 inclusive, and whose parents/guardians had given consent, were eligible for the study. Of 297 families who had initially assented to the study, 222 subsequently provided a written informed consent; 14 were not eligible, 44 had decided against participating and 17 were lost to follow up.

2.1.3. Data and sample collection

An initial birth interview was undertaken using a researcher-administered custom questionnaire which included details of demographics, pregnancy and initial feeding. Additionally, maternal and neonatal notes, prescription charts and, where used, anaesthetic charts were reviewed for antibiotic use and labour and delivery details. We visited all participants at 6 weeks, 6, 9, 12, 18 and 24 months of age. At home visits, the infant's posterior oropharyngeal wall was swabbed with a double-headed nylon flocked swab (Copan diagnostics) (using a tongue depressor to avoid contamination from surrounding structures) by the research fellow or research nurse, who was trained in the technique. Oropharyngeal swabs were returned to the laboratory at room temperature and frozen at -80 °C within 10 h of collection. An interviewer-administered custom questionnaire was conducted at each home visit. The interviewer was unaware of any microbiota findings at the time of home visits and medical notes review.

2.2. Data sources and definition of outcomes

After the 2 years visit, medical notes were requested from the general practitioner (GP). The completeness of the data obtained was reviewed and where necessary the GP was contacted for further information or the previous GP was contacted. We transcribed primary care health care records, which include diagnoses, prescription of medications (including antibiotics), records of Emergency Room or out-of-hours visits, inpatient discharge letters and outpatient appointment letters. If these were incomplete and not available from the GP, where possible these documents were obtained from the source hospital. GP notes were reviewed for all consultation details; antibiotic and other medication use, immunisations and wheeze diagnoses were recorded.

Doctor-diagnosed wheeze was defined from the medical records as a recorded diagnosis of wheeze, auscultation-confirmed wheeze, or documentation by the physician of a history consistent with wheeze and a prescription of bronchodilator.

Recurrent wheeze was defined as more than one episode of doctor-diagnosed wheeze.

Antibiotic usage was ascertained both according to parental report and prescription data in the 4 weeks prior to the visit.

2.3. Assessment of microbiome

2.3.1. Bacterial DNA extraction and sequencing

DNA was extracted within 10 days of collection using the FastDNA Spin kit for soil (MPBiomedicals). The final elution step was into Tris (10 mM) low- ethylenediaminetetraacetic acid (0·1 mM) buffer; otherwise all steps were according to the manufacturer's protocol. An amplicon library was created by polymerase chain reaction (PCR)-amplifying the V3-V5 region of the 16S rRNA gene using a primer pair tagged with individually unique 12 base pair error correcting Golay barcodes [14]. PCR was performed in quadruplicate as described previously [14] with 2 changes: an increased volume of template solution was used (2 μl, total reaction volume still 25 μl), and the reaction was cycled 35 times. Amplicons were pooled, purified, size selected, and then pooled to an equimolar concentration and submitted for 454 pyrosequencing using the GS-FLX-Titanium (Roche) with the Roche Amplicon Lib-L protocol. Negative and positive controls (samples repeated across runs) were included to assess contamination and inter-run variability.

The amplicons were quality filtered and denoised using denoise_wrapper.py (denoising software for multiple 454 runs) within Quantitative Insights Into Microbial Ecology (QIIME) pipeline version 1·9 [15]. Chimeras were removed using the usearch61 algorithm in QIIME. Sequences were clustered to 97% similarity using uclust, using the SILVA rRNA database (SSU_REF119) as reference [16]. Singletons and amplicons present in only a single sample were removed. Sequencing data is available at the European Nucleotide Archive: PRJEB6349.

2.3.2. qPCR method

Quantification of bacterial 16S rRNA-encoding genes was performed by real-time PCR with TaqMan hydrolysis probes on a StepOne™ Real-Time PCR System (Applied Biosystems). The primers used, targeting the V3 and V4 regions of the 16S rRNA-encoding gene, were 5′ ACTC CTACGGGAGGCAG 3′ (forward) and 5′ GACTACCAGGGTATCTAATCC 3′ (reverse), and the probe used was 5′-(FAM)-TGCCAGCAGCCGCGGT AATAC-(BHQ-1)-3′ (where FAM is 6-carboxyfluorescein and BHQ1 is Black Hole Quencher 1 [17]. After an initial denaturation for 10 min at 95 °C, 40 cycles of amplification for 15 s at 95 °C, and 60 s at 60 °C were performed. Final cooling was performed at 4 °C.

Quantification of 16S rRNA-encoding genes was achieved by comparison with standard curves for known 16S rRNA concentrations. These were created by cloning amplified *Pseudomonas aeruginosa* PAO1 16S rRNA gene into TOPO TA Cloning vector. Qiagen plasmid extraction kit was used to obtain plasmid DNA, gel purified, and quantified using picogreen. Plasmid DNA was diluted to a working concentration of 1E+09 molecules/μl. Dilutions ranging from 1E+04 to 1E+08 molecules/μl were used on each qPCR plate to create standard curves.

2.3.3. Data analysis

Bacterial density was compared using the Friedmann test with a post-hoc Nemenyi test to assess differences between time points. To mitigate large differences in library sizes, rarefaction was used prior to alpha diversity analysis and the cumulative sum scaling (CSS) normalization was applied to OTU raw counts for all other analysis [18]. The variability within bacterial communities (alpha diversity) was assessed with the Shannon and Simpson indices obtained using *phyloseq* R package [19] and, consequently, were compared using Friedmann test with post-hoc Nemenyi tests. The differences among bacterial communities (beta diversity) were assessed by Bray-Curtis distance and plotted using principal coordinates analyses (PCoA). The effect of time and disease state was tested using the *adonis* procedures implemented in the *vegan* R package [20].

To ascertain the association between microbiota and clinical outcomes, we used a recently developed method for longitudinal marker-gene surveys based on smoothing splines [21], which uses longitudinal profiling of the microbiota communities to explore the complex relationships between community dynamics and phenotypes of interest. Specifically, this method detects time intervals for which bacteria are differentially abundant. Given observations at multiple time points, the difference in abundance across time is modelled as a function. Then, using group membership permutations, a null distribution of areas under the difference curve is estimated and used to detect time intervals of differential abundance [22]. In the longitudinal differential abundance analysis, to guarantee detection of relevant differences, we considered time intervals to be of potential significance if the absolute difference between two groups was above 0·2 [21]. All models were

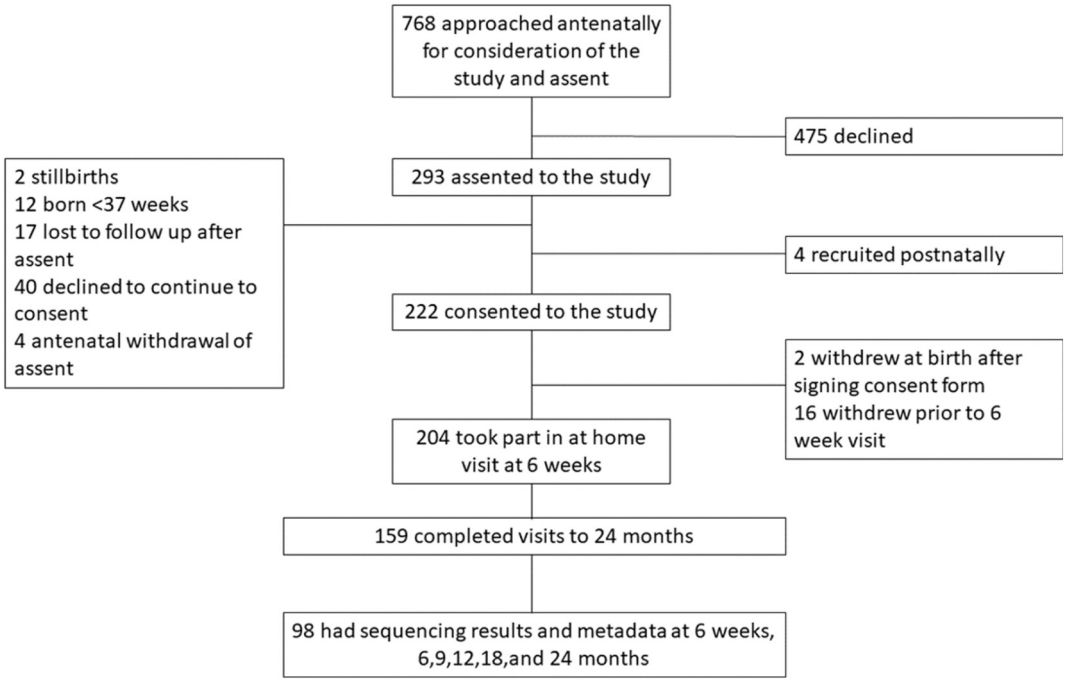

Fig. 1. Consort diagram showing the participant flow through the DORMICe study.

adjusted for ethnicity, family history of atopy (fixed), presence of fever and the use of antibiotics in the 4 weeks prior to visit (time-varying) (for further details refer to the online supplement). To filter out noise, Operational Taxonomic Units (OTUs) (cluster of reads with ≥97% similarity of the 16S rRNA gene) found to be present less than three times in at least 20% of samples over time were removed from the analysis [22]. Analyses were implemented using the R package *metagenomeSeq* [23] using the *fitTimeseries* function.

3. Results

3.1. Characteristics of the study population

Participant flow is shown in Fig. 1; 768 prospective parents were approached, of whom 297 (39%) assented for their baby to participate once born. Children were born between January and October 2013; 204/297 (69%) parents consented to follow-up and completed at least one home visit, and 159/204 (78%) completed the 24-months visit. Children for whom sequencing and outcome data were not available for all time points were excluded ($n = 61$). There were no significant differences in demographic characteristics and outcomes between children included and excluded from this analysis, except for ethnicity (Table 1).

Of 98 children with complete data, 25 were not prescribed antibiotics by age 2 years, 48 received 1–2 courses, 21 had 3–4 courses, and 4 were prescribed ≥5 courses. Doctor-diagnosed wheeze was confirmed in 26 children (median age of onset [interquartile range] 299 days [230–567]), of whom 11 had recurrent wheezing. Table 2 shows demographic characteristics and environmental exposures among children with and without doctor-diagnosed wheeze. Wheeze was significantly more common among children with a family history of atopy, those who had fever in the 4 weeks prior to the 9-month visit, and those who received antibiotics in the 4 weeks prior to the 12-month visit (Table 3).

3.2. Bacterial diversity and density

We obtained 1,559,882 sequencing reads from 588 swabs from 98 children, with a mean number of 2653 reads per sample. Fig. 2 shows the bacterial density by time point, represented by the quantity of 16S DNA recovered by qPCR. The bacterial density increased significantly in the first 9 months ($p < 0{\cdot}001$), and then remained unchanged thereafter.

Taxonomy of the 40 filtered OTUs (present more than three times in at least 20% of samples over time) is shown in Table E1. Fig. 3 shows their relative abundance at each time point. We observed a change in the composition of the oropharyngeal microbiota over the first 2 years of life (Fig. 3). The oropharyngeal microbiota was consistently characterized by high prevalence of the genus *Streptococcus*. Both alpha diversity measures show an increasing trend over 2 years (Fig. 4). Significant differences in alpha diversity were found throughout the entire period and were particularly striking when comparing 6-weeks samples with all others (Table E2). Considering Bray–Curtis dissimilarities, PCoA analysis (Fig. 5A&B) shows a difference in the beta diversity between 6 weeks and the other time point samples, highlighting the progression of the microbiota composition during the first 2 years of life (PERMANOVA analysis; p-value < 0·001).

3.3. Oropharyngeal microbiota in children with and without doctor-confirmed wheeze

Fig. 6 shows the relative abundances of the 40 filtered OTUs among children with and without doctor-confirmed wheeze. There was a similar progression over 24 months in terms of the OTUs present, although we noted an increase in the level of *Neisseria* at age 18 months among children who wheezed. Considering Bray–Curtis dissimilarities, no differences in microbiota composition were found between wheezers and non-wheezers (PERMANOVA analysis, $p = 0{\cdot}690$, Fig. 5C&D).

Results of the longitudinal differential abundance analysis are presented in Table 4. After adjusting for possible confounders (ethnicity, family history of atopy, presence of fever and the use of antibiotics in the 4 weeks prior to visit) and after Bonferroni correction for multiple testing, three OTUs from the genera *Neisseria, Prevotella* and *Granulicatella* were found to be differentially abundant between children with and without doctor-diagnosed wheeze. There was a significant and substantial increase in the abundance of a *Neisseria* OTU over time in children with wheeze ($p = 0{\cdot}003$), while in those without we observed a significant increment in the abundance of a *Granulicatella* OTU ($p = 0{\cdot}012$) and of a *Prevotella* OTU ($p = 0{\cdot}018$). These three OTUs were equally abundant among wheezers and non-wheezers at

Table 1
Characteristics of the study population presented for participants with complete and incomplete data for all time points.

Characteristics		Complete data ($N = 98$)	Incomplete data ($N = 106$)[a]	p-Value
Gender	Male	48 (49%)	49 (46%)	0·779
Ethnicity	*Asian British*	6 (6%)	14 (13%)	**0·011**
	Black British	4 (4%)	16 (15%)	
	Mixed	25 (26%)	16 (15%)	
	Other	10 (10%)	13 (12%)	
	White	53 (54%)	47 (44%)	
Season of birth	*Spring*	32 (33%)	38 (36%)	0·502
	Summer	34 (35%)	38 (36%)	
	Autumn	4 (4%)	8 (8%)	
	Winter	28 (29%)	22 (21%)	
Gestation (days)		282·5 [273·2; 289·0]	283·0 [277·0; 288·0]	0·849
Birth weight (gr.)		3500 [3130; 3810]	3535 [3140; 3805]	0·739
Family history atopy	*Yes*	27 (28%)	22 (21%)	0·325
Mode of delivery	*Vaginal birth*	67 (68%)	82 (77%)	0·159
Neonatal antibiotics	*Yes*	9 (9%)	8 (8%)	0·801
Intrapartum antibiotics	*Yes*	14 (14%)	11 (11%)	0·404
Smoking at home	*Yes*	27 (28%)	25 (24%)	0·525
Furry pet at home	*Yes*	12 (12%)	15 (14%)	0·837
Respiratory symptoms a week either side of visit	*6 weeks*	18 (19%)	20 (19%)	1·000
	6 months	33 (34%)	26 (31%)	0·752
	9 months	41 (42%)	27 (34%)	0·282
	12 months	36 (37%)	23 (32%)	0·625
	18 months	52 (53%)	29 (43%)	0·200
	24 months	32 (33%)	21 (34%)	0·863
Wheeze a week either side of visit	*6 weeks*	0 (0%)	1 (1%)	1·000
	6 months	3 (3%)	0 (0%)	0·253
	9 months	3 (3%)	0 (0%)	0.253
	12 months	1 (1%)	0 (0%)	1·000
	18 months	1 (1%)	0 (0%)	1·000
	24 months	1 (1%)	0 (0%)	1·000
Fever a week either side of visit	*6 weeks*	4 (4%)	2 (2%)	0·430
	6 months	11 (11%)	9 (11%)	1·000
	9 months	15 (15%)	5 (6%)	0·093
	12 months	10 (10%)	7 (10%)	1·000
	18 months	9 (9%)	8 (12%)	0·608
	24 months	13 (13%)	6 (10%)	0·620
Antibiotics in the 4 weeks prior to visit	*6 weeks*	8 (8%)	3 (3%)	0·123
	6 months	3 (3%)	4 (5%)	0·705
	9 months	11 (11%)	6 (8%)	0·452
	12 months	11 (11%)	4 (6%)	0·277
	18 months	8 (8%)	4 (6%)	0·763
	24 months	7 (7%)	4 (7%)	1·000

[a] Sample size changes according to visit attendance. The reference denominator is $N = 106$, which is referred to the participants who entered the study by attending visit at 6 weeks. For the time-varying characteristics, denominators are $N_{6w} = 106$, $N_{6m} = 84$, $N_9 = 80$, $N_{12} = 71$, $N_{18} = 67$, $N_{24} = 61$. Wheeze a week either side of visit is by parental report in this table only (due to the incomplete data with some participants missing GP notes).

Table 2
Characteristics of the study population presented for wheezers and non-wheezers.

Characteristics		Wheeze ($N_w = 26$)	Non-wheeze ($N_{nw} = 72$)	p-Value
Gender	Male	16 (62%)	32 (44%)	0·172
Ethnicity	*Asian British*	1 (4%)	5 (7%)	0·935
	Black British	1 (4%)	3 (4%)	
	Mixed	8 (31%)	17 (24%)	
	Other	2 (8%)	8 (11%)	
	White	14 (54%)	39 (54%)	
Season of birth	*Spring*	9 (35%)	23 (32%)	0·404
	Summer	7 (27%)	27 (38%)	
	Autumn	0 (0%)	4 (6%)	
	Winter	10 (38%)	18 (25%)	
Gestation (days)		278·5 [272;289·0]	283·5 [274·0;290·0]	0·346
Birth weight (gr.)		3520 [3152;3718]	3460 [3135;3815]	0·812
Number of months breastfed		6·5 [2·0; 9·5]	6·0 [3·0; 10·7]	0·938
Family history atopy	*Yes*	12 (46%)	15 (21%)	**0·020**
Mode of delivery	*Vaginal birth*	17 (65%)	50 (69%)	0·806
Neonatal antibiotics	*Yes*	0 (0%)	9 (13%)	0·107
Intrapartum antibiotics	*Yes*	1 (4%)	13 (18%)	0·104
Smoking at home	*Yes*	7 (27%)	20 (28%)	1·000
Furry pet at home	*Yes*	2 (8%)	10 (14%)	0·507
Respiratory symptoms a week either side of visit	*6 weeks*	6 (23%)	12 (17%)	0·558
	6 months	11 (42%)	22 (31%)	0·335
	9 months	14 (54%)	27 (38%)	0·169
	12 months	11 (42%)	25 (35%)	0·489
	18 months	11 (42%)	41 (57%)	0·253
	24 months	9 (35%)	23 (32%)	0·811
Fever a week either side of visit	*6 weeks*	1 (4%)	3 (4%)	1·000
	6 months	4 (15%)	7 (10%)	0·475
	9 months	8 (31%)	7 (10%)	**0·022**
	12 months	2 (8%)	8 (11%)	1·000
	18 months	1 (4%)	8 (11%)	0·438
	24 months	4 (15%)	9 (13%)	0·740

Square brackets indicate interquartile range around the median. Fisher's Exact and Mann-Whitney test where appropriate. Bold indicates p value < 0.05.

Table 3
Antibiotic usage according to parental report and prescription data in the 4 weeks prior to the visit.

Antibiotic variable[a]	*Visit*	*Wheezers N = 26*	*Non-wheezers N = 72*	*p-values*
Antibiotics in the 4 weeks prior to visit	*6 weeks*	1 (4%)	7 (10%)	0·677
	6 months	1 (4%)	2 (3%)	1·000
	9 months	5 (19%)	6 (8%)	0·154
	12 months	8 (31%)	3 (4%)	**0·001**
	18 months	1 (4%)	7 (10%)	0·677
	24 months	1 (4%)	6 (8%)	0·671
Combined antibiotics within a month medical and parental report	*6 weeks*	1 (4%)	7 (10%)	0·677
	6 months	1 (4%)	2 (3%)	1·000
	9 months	7 (27%)	6 (8%)	**0·037**
	12 months	8 (31%)	3 (4%)	**0·001**
	18 months	1 (4%)	7 (10%)	0·677
	24 months	3 (12%)	6 (8%)	0·696
Antibiotics prescription within month of visit	*6 weeks*	1 (4%)	7 (10%)	0·677
	6 months	1 (4%)	1 (1%)	1·000
	9 months	7 (27%)	3 (4%)	**0·003**
	12 months	7 (27%)	1 (1%)	<**0·001**
	18 months	1 (4%)	3 (4%)	1·000
	24 months	2 (8%)	4 (6%)	0·654

The combined variable represents the antibiotic usage according to both parental report and prescription data. Fisher's Exact was used to assess differences in proportions. Bold indicates p value < 0.05.

[a] There are three different variables for antibiotic use each with its own merits. The variable antibiotics in the 4 weeks prior to visit is according to parental report. This variable will thus not include situations where the GP has prescribed antibiotics but these are not subsequently given. For example, the case of a delayed script for an upper respiratory tract infection in case of persistence of symptoms. It will include antibiotics given by other providers for example dentists, or out of hour services which may not have been included in the GP notes or antibiotics given abroad. The antibiotics prescription within a month of visit was derived from prescriptions within the GP notes. The combined antibiotics within a month includes both the courses reported by the parent and those from the GP notes and is the most comprehensive account of antibiotic use.

6 weeks of age; the *Neisseria* OTU became differentially abundant over the period between 9 and 24 months, with an increasing trend throughout the entire period, while the estimated difference function for the *Granulicatella* OTU showed that this genus was differentially abundant between 9 and 12 months, and the *Prevotella* OTU became differentially abundant after the 18th month (Fig. 7). Significant changes in both *Neisseria* and *Granulicatella* OTUs were detected in the period preceding the median age of onset of wheeze (299 days).

To evaluate the robustness of the results with respect to the outcome definition, we perfomed a sensitivity analysis comparing longitudinal profiles of microbiota of children with recurrent wheezing (two or more occurrences of doctor diagnosed wheeze) and those who had never wheezed. After Bonferroni correction, three OTUs had significantly lower abundance among recurrent wheezers: a *Granulicatella* OTU and two *Prevotella* OTUs (Table 5). The shift in the *Granulicatella* OTU abundance appeared to last a longer period than for the wheeze group as a whole, being detected from 9 months and lasting until 24 months. The OTUs of the *Prevotella* genus were detected at a later time-point; OTU76 was differentially abundant from 12 to 24 months, and OTU43 from 12 months to 24 months. While changes in the *Granulicatella* OTU were significant shortly before the median age of onset of wheeze, the other two OTUs became differentially abundant after the median age of onset (Fig. E1).

4. Discussion

Our data suggest that there is a time window which opens before age one year in which colonization of the oropharynx with *Neisseria* is positively, and with *Granulicatella* species negatively, associated with doctor-diagnosed wheezing by age two years. Although we cannot infer causality, we note that some of the identified shifts in OTUs occurred prior to the development of wheeze, whilst other shifts occurred following this event.

To the best of our knowledge, our study is the first detailed longitudinal evaluation of the oropharyngeal microbiota during early life in an unselected sample of healthy infants from the community, and its relationship with wheeze. These results suggest that changes in the oropharyngeal microbiota development may modulate wheeze development and might inform the design of preventive treatments.

There are limitations to our study. Through our filtering procedure, some potentially important OTUs may have been excluded. However, the filtering process was necessary to moderate the effect of measurement errors and noise. Another limitation arises from the relatively small sample size, thus the study may have been underpowered to detect all associations of the oropharyngeal microbiota with wheeze. We acknowledge that some of the negative findings may be secondary to low statistical power of the study. In addition, for our analysis we required sequencing data for every timepoint with the outcome data, and 98/159 (62%) of the participants completing the study had this available, and 159/220 (72%) of those whom initially consented completed the study. There is thus a possibility of selection bias as a result of this loss to follow up. Whilst there were minimal differences between baseline characteristics of those included in the analysis compared to those either without sufficient data or who did not complete follow up, sample sizes were small and there may not have been sufficient statistical power to detect differences. Further external validations are needed to ascertain generalizability of our findings.

There are other potential confounders which we have not accounted for, including viral infection. There are known interactions between the upper respiratory tract microbiota and respiratory viruses [24], (including early RSV infection which is also associated with later wheeze)

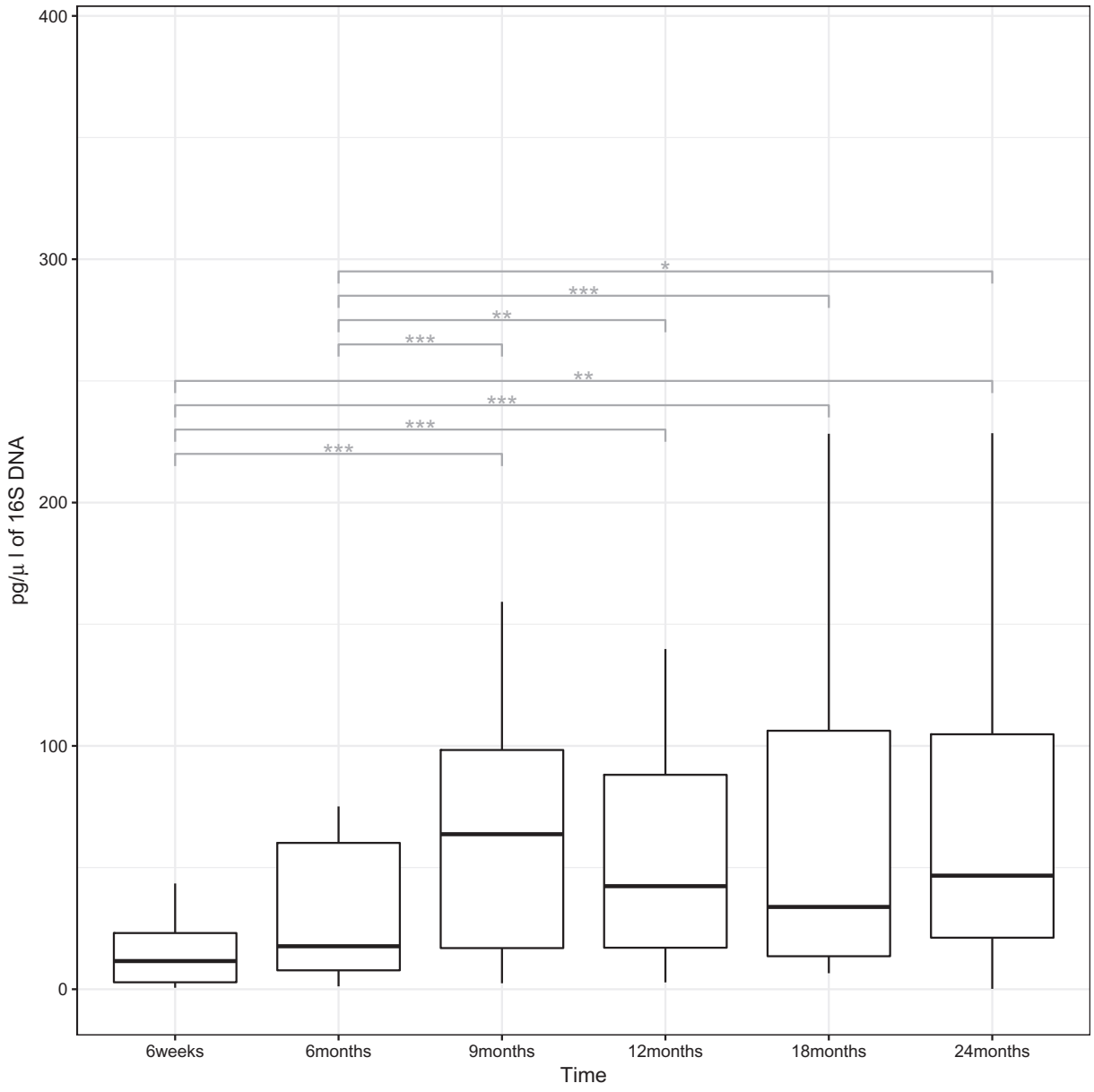

Fig. 2. Boxplot of qPCR data. Outliers have been removed to visualize the overall distribution. The upper and lower hinges correspond to the 75th and 25th percentile and the line across to the median; the whiskers extend to 1.5x interquartile range (IQR) from the hinge. A statistically significant change in bacterial density is denoted by representing $p < 0{\cdot}05$, $^{**}p < 0{\cdot}01$ and $^{***}p < 0{\cdot}001$. Significant differences were assessed through post-hoc Nemenyi tests.

which may have confounded our analysis. Whilst reviewing GP records we made note of any hospital admissions and among the 98 participants studied, two had been admitted with bronchiolitis.

Using microbial marker genes such as the 16S rRNA gene have their limitations particularly when attempting to discriminate between species [25]. We cannot comment further on the role of individual species, and this is particularly an issue within the *Streptococcus* genus, limiting the comparison of our findings to previous culture-based studies [8].

We endeavoured to visit children when they were well, but 36.5% had symptoms of a mild respiratory illness, 3.8% had wheezed and 10.6% had fever within the week preceding or following the visit. We adjusted analyses for the presence of recent fever. We have included a health professional's interpretation of report of wheeze with a prescription of bronchodilator for wheeze – this may have led to an overdiagnosis of wheeze compared to solely including where the notes included auscultation of wheeze. However, some of the medical notes were brief and did not include details of auscultation thus we may have missed cases if we had not used this definition.

Lastly, we present wheeze data to age two years, after which wheezing patterns continue to develop [26]. Therefore, further follow-up will be necessary to draw more definitive conclusions. Assessment of allergic sensitization and lung function, which we did not conduct, would be useful at follow-up.

Our study has several strengths; we have used a community-based unselected cohort, which increases its relevance to the general population. We have meticulously collected clinical data from primary care records, so that our definition of wheeze is not reliant on parental report which has its limitations [27], but was determined by a health professional. We collected detailed data on potential confounders. We endeavoured to comprehensively explore antibiotic use, which is widespread in young children and may introduce potentially major bias, at least temporarily perturbing the microbial community structure [28], and have adjusted our analyses for the recent use of antibiotics.

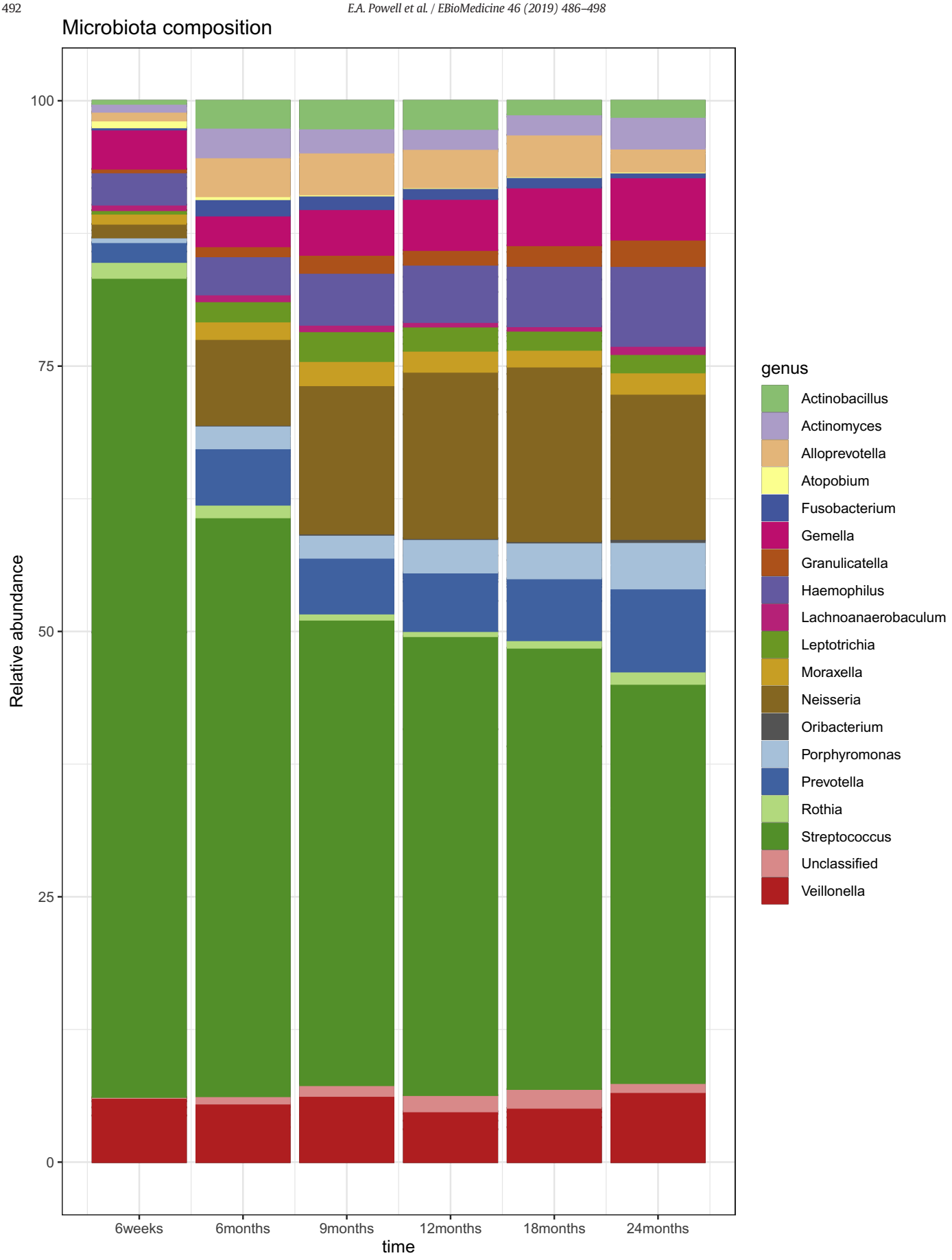

Fig. 3. Composition of the microbiota from 6 weeks to 24 months of age. Relative abundances from the raw data for the 40 most abundant OTUs at the genus level at each time point. Figure obtained using the R package phyloseq [19].

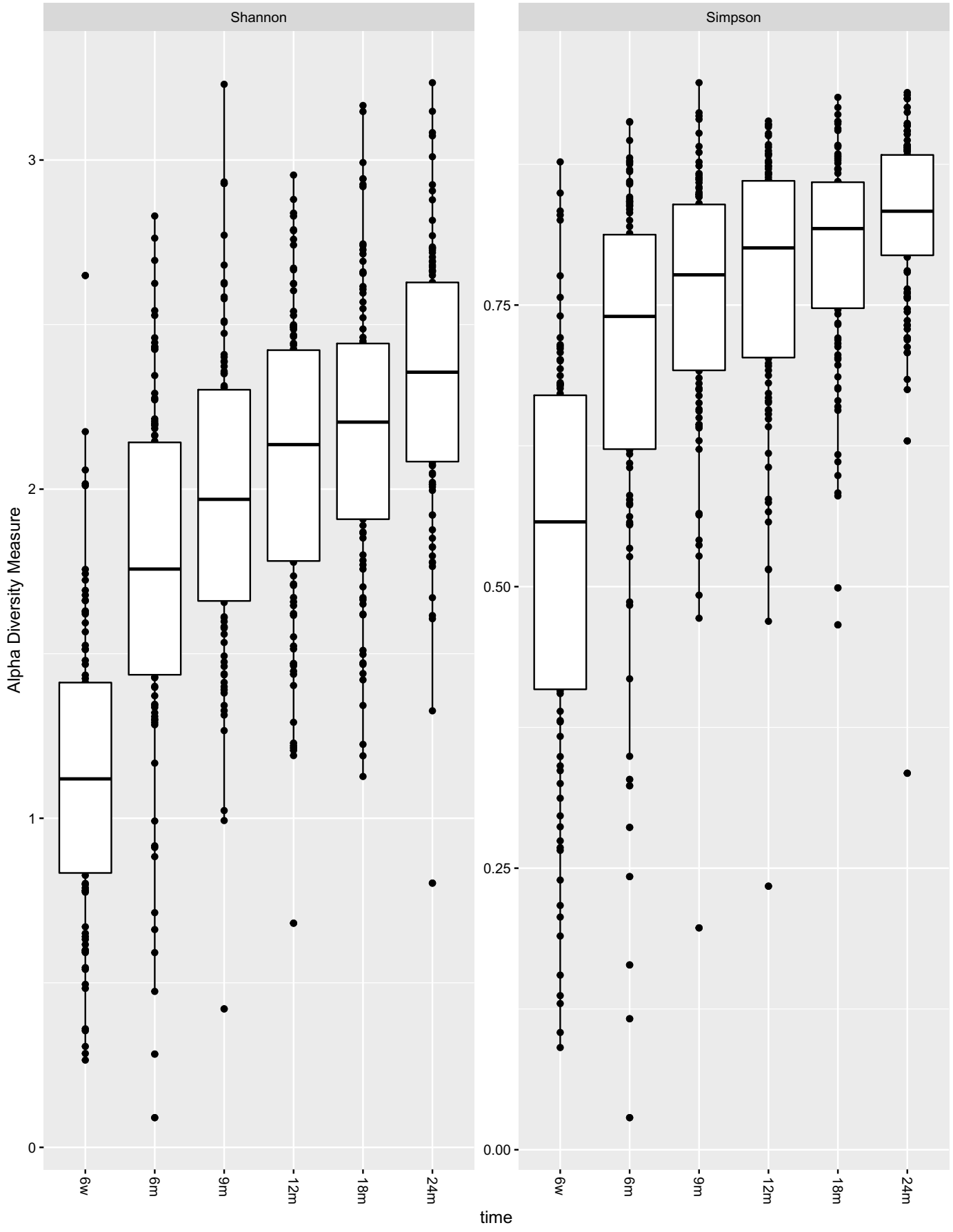

Fig. 4. Comparison of microbiome alpha diversity across time. To guarantee reliability of the estimates, OTUs that were not present in at least one sample were removed from this analysis and rarefaction was used to normalise raw count data. Significant differences are assessed through Friedmann test with post-hoc Nemenyi tests (in table E2). Figure obtained using the R package phyloseq [19].

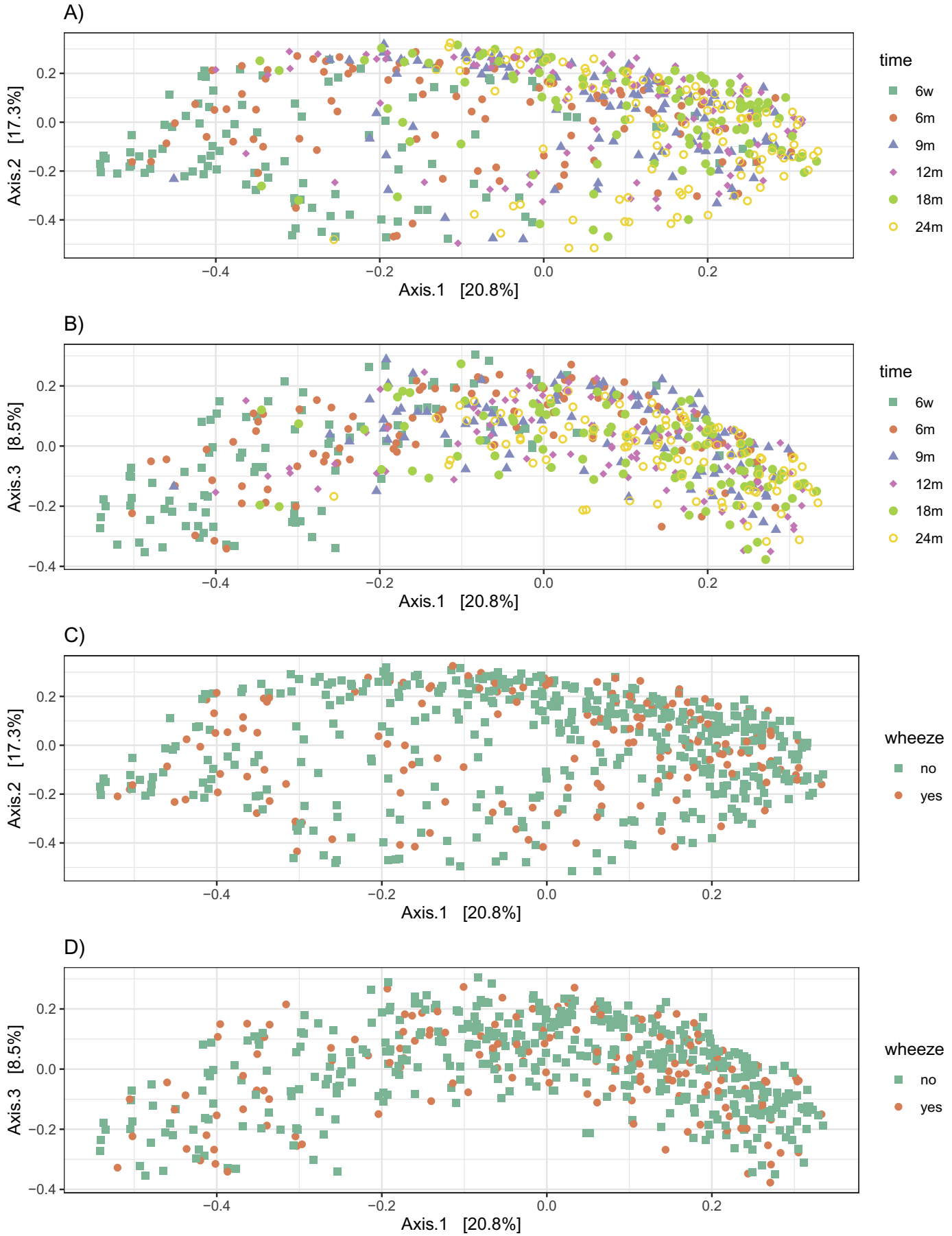

Fig. 5. Analysis of beta diversity. Principal coordinates analysis (PCoA) derived from Bray-Curtis distance ($p < 0{\cdot}001$ by adonis). Colours represent the different time points for the upper two plots (A&B) and wheeze diagnosis for the lower two plots (C&D). For each axis, in square brackets, the percent of variation explained was reported. Three-dimensional solution is retained and plots for dimension 1 *versus* dimension 2 and for dimension 1 *versus* dimension 3 are shown.

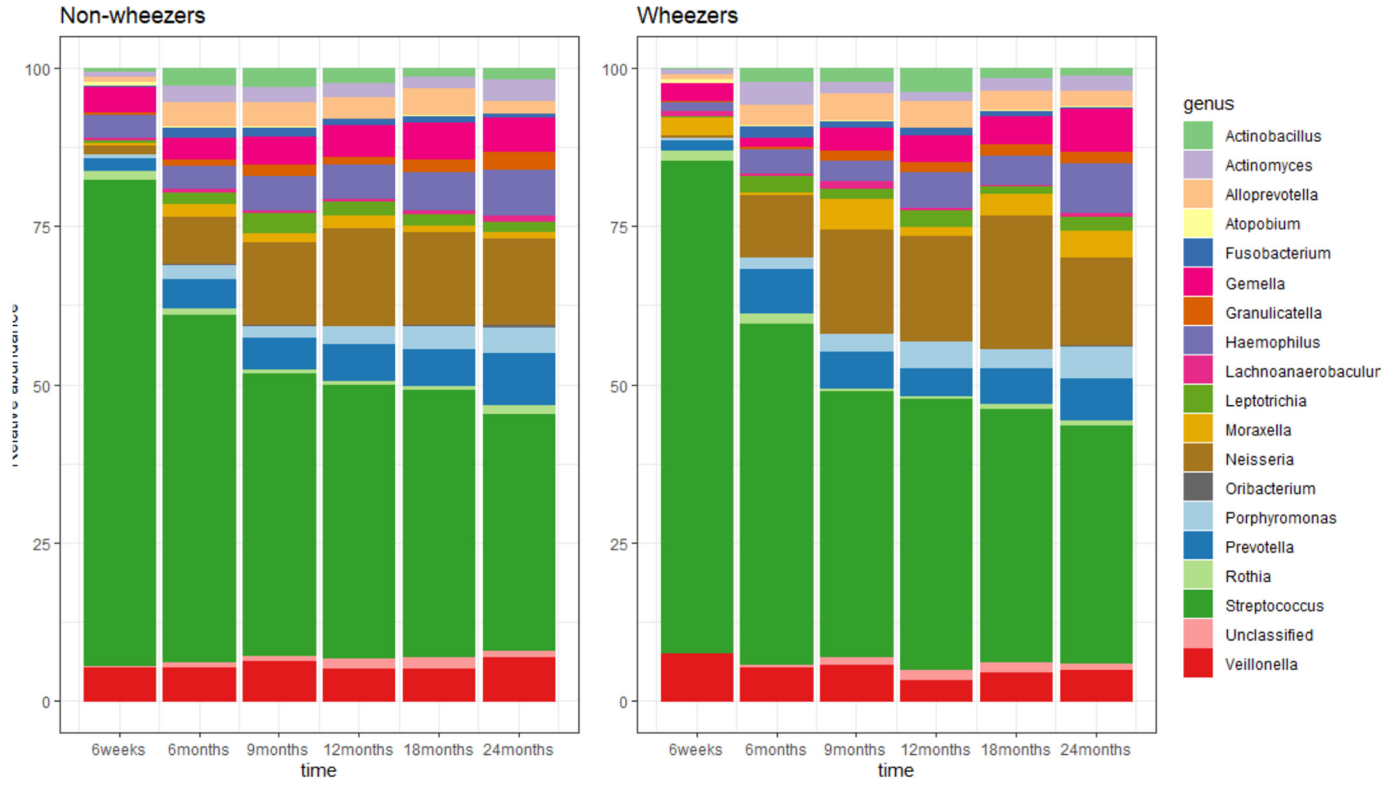

Fig. 6. Composition of the microbiota from 6 weeks to 24 months of age stratified by wheeze status. Relative abundances from the raw data for the 40 most abundant OTUs (aggregated by genus) at each time point in wheezers and non-wheezers. Figure obtained using the R package phyloseq [19].

We have used a non-culture method to identify the majority of the microbiota, avoiding the bias introduced by culture. In recognition of the importance of the development of the microbiota over time, we have used a recently developed method of analysis to take a longitudinal view of the data.

In our study, the respiratory microbiota through the first two years of life was dominated by *Streptococci*, which is consistent with previous studies [29,30]. We observed an increase in diversity of the oropharyngeal microbiota with increasing age, with a changing pattern of OTUs, and a fairly smooth temporal change over time. In contrast to the nasopharyngeal microbiota [31], the oropharyngeal microbiota appears less varied and is dominated by a single genus. This topography of the microbiota may affect the relationship between the microbiota at different niches and preschool wheezing.

An association between wheeze, asthma and the resident microbiota has been suggested by epidemiological studies showing a reduction in the incidence of these conditions in children who are exposed to a richer environmental microbiota through growing up on a farm, and/or by their mothers working in stables when pregnant [32]. These studies were followed by an increasing body of evidence for an association between the composition of the early respiratory tract microbiota and subsequent respiratory health, including its relation to wheeze [9,11], asthma [10], bronchiolitis [33], pneumonia [33], and upper respiratory tract infections [31,34]. Using longitudinal analyses, we questioned whether a change over time in specific species were associated with doctor-diagnosed wheeze. We have found a positive temporal association between the presence of a *Neisseria* OTU and doctor-diagnosed wheeze, and this difference was observed from age 9 months. Although we cannot ascertain the causal relation of changes in the microbiota and wheeze development, we have shown that the differential abundance of the OTUs started in the period preceding the median age of onset. Conversely, a *Granulicatella* OTU was negatively associated with wheeze. We have found a negative association between recurrent wheeze and two *Prevotella* OTUs and a *Granulicatella* OTU, albeit with small numbers of recurrent wheezers.

Similar features with higher prevalence of *Neisseria* in wheezing infants compared to non-wheezers have previously been demonstrated in a cross-sectional study in 10-month old infants in Ecuador [35]. However, there were also differences to our study, with *Prevotella* being associated with wheeze in this case-control study. *Veillonella* was observed more frequently in controls. Other cross-sectional studies have associated the phylum *Proteobacteria* with airway disease [5], and have shown a difference in the nasal microbiota in terms of both alpha and beta diversity and the abundance of a *Moraxella* OTU [12]. School-age asthmatic children had a higher abundance of *Moraxella*, but only if they had not grown up on a farm,

Table 4
Results of longitudinal differential abundance analysis in early-life wheeze.

Differential abundance analysis: Children with doctor-confirmed wheeze (n = 26) *vs.* no wheeze (n = 72)					
Genus	Interval start	Interval end	Differential area	p-value	Adjusted p-value
OTU42: *Granulicatella*	9 months	12 months	(−) 1·118	0·004	**0·012**
OTU43: *Prevotella*	18 months	24 months	(−) 1·147	0·006	**0·018**
OTU68: *Neisseria*	9 months	24 months	(+) 5·993	0·001	**0·003**

Time intervals are considered of interest if the absolute difference between two groups is above 0·2. The sign of the difference area provides information on the direction of the abundance shift. OTUs are considered differentially abundant if their adjusted p-values < 0·05.

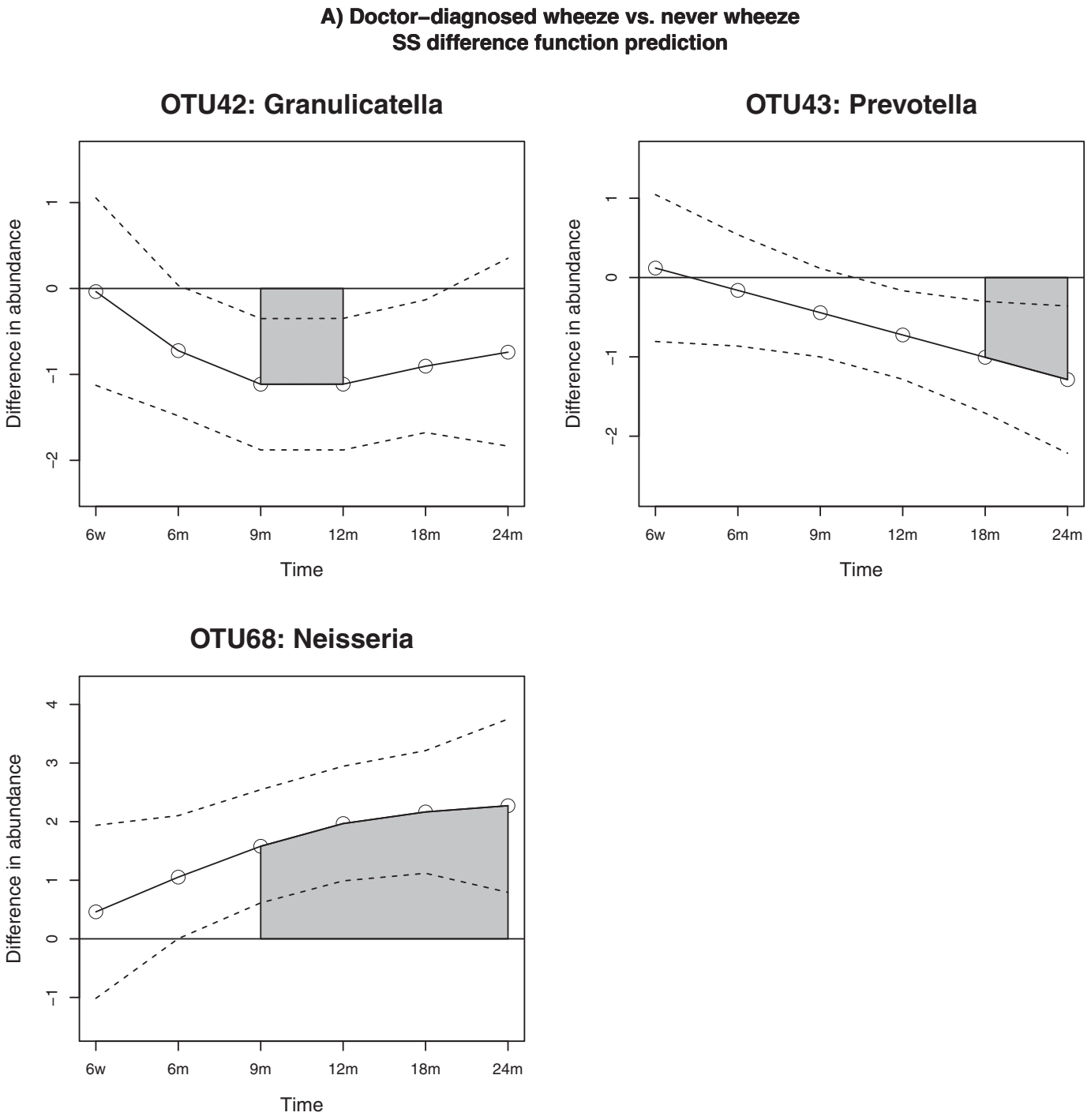

Fig. 7. Relationship between microbiota community dynamics and the development of wheeze assessed using longitudinal differential abundance analysis adjusted for ethnicity, family history of atopy, presence of fever and recent use of antibiotics. The grey-shaded areas show the time interval for which the OTU was differentially abundant, the dashed lines show the Bayesian confidence intervals, and the solid line the estimated difference function.

whereas in farm children the same *Moraxella* OTU was not associated with asthma, suggesting the farm environment had a modifying effect [12]. In this study, the authors found an association at school age between the nasal microbiota and asthma but no such association between the microbiota of the throat and asthma. It may be that association between the microbiota and clinical phenotype are different prior to disease onset than in already established disease, and there may be different associations with disease, with the microbiota of the young infant where the microbiota is still developing and that of the established microbiota. The site of swabbing and therefore the niche sampled in the study by Depner *et al*, was also different to our study – they swabbed the soft palate and tonsils, whereas we took swabs from the oropharynx.

Granulicatella is a well-recognised commensal of the upper respiratory tract. Its presence in the microbiota has been negatively associated with the presence of *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Moraxella catarrhalis,* an association modulated by antibiotic use [28]. In the same study, *Granulicatella*, *Rothia*, *Gemella*, *Actinomyces and Veillonella* were correlated taxa which as a group were less likely to be associated with *S. pneumoniae* carriage. Our

Table 5
Sensitivity analysis for outcome definition.

Differential abundance analysis: Children with recurrent wheeze ($n = 11$) *versus* no wheeze ($n = 72$).					
Genus	Interval start	Interval end	Differential area	p-value	Adjusted p-value
OTU42: Granulicatella	9 months	24 months	(−)4·999	0·001	**0·004**
OTU43: Prevotella	18 months	24 months	(−)2·637	0·007	**0·028**
OTU68: Neisseria	12 months	24 months	(+)3·581	0·015	0·060
OTU76: Prevotella	12 months	24 months	(−)2·090	0·005	**0·020**

Time intervals are considered of interest if the absolute difference between two groups is above 0·2. The sign of the difference area provides information on the direction of the abundance shift. OTUs are considered differentially abundant if their adjusted p-values < 0·05. Bold indicates p value < 0.05.

finding of a negative association of *Granulicatella* with wheeze may reflect underlying beneficial polymicrobial interactions which are protective against the development of wheeze. Interestingly, a recent family-based cross-sectional study found a higher abundance of *Granulicatella* in the sputum of the non-asthmatic sibling compared to their asthmatic sibling [36]. Rosas-Salazar *et al*. [37], followed a cohort of infants who had had RSV infection and monitored for subsequent development of wheeze. They found an increase in detection and abundance of *Lactobacillus* in the nasopharynx at the time of RSV infection in those who did not develop wheeze compared to those who did. Whilst their cohort had different characteristics (all had had RSV infection, the definition of wheeze was reliant on parental report alone, and the analysis was at one timepoint point only), the results lead to a similar hypothesis of a potential benefit of commensal bacteria protecting against wheeze development.

We did not find the association between early *Streptococcus* and wheeze seen by others, which may be due to methodological differences; Bisgaard et al. [8], used a culture method to detect *S. pneumoniae*, and Teo et al. [9], dichotomised the *Streptococcus* load in early nasopharynx samples analysed using next generation sequencing. In addition to variation by anatomical region sampled, results from next generation sequencing studies can also vary by sequencing platform used (*e.g.* Illumina used by Teo et al. [9], *versus* Roche 454 platform which we used), and the region of 16S rRNA sequenced (V4 *versus* V3-V5 respectively).

We observed changes in the microbiota occurring later in infancy than we had anticipated based on previous evidence. It is unclear whether the changes in microbiota described in our study are a result of a disordered or conversely beneficial ecosystem, and whether the microbiota have had a role in the airway and immune system development resulting in or protecting from wheeze.

In conclusion, we identified a window in the development of the oropharyngeal microbiota between 9 and 24 months of age where, for most of the OTUs, shifts in abundance occurred. The abundance of OTUs from the genera *Neisseria* and *Granulicatella* was different according to whether the infant was diagnosed with wheeze. Future work will test whether this relationship holds for asthma at school age.

Acknowledgements

We would like to thank the DORMICe research team including research nurses who worked on the project, and most importantly the families who took part in the study.

Funding sources

Research funding from the Winnicott Foundation, Meningitis Now and an unrestricted research grant from Micropathology Ltd. Meningitis Now and the Winnicott Foundation had no role in the study design, analysis, interpretation or report writing. *CF* co-founder of Micropathology Ltd. is an academic supervisor of EP, and was involved in study design and report editing.

Declarations of Competing Interest

Dr. Custovic reports personal fees from Novartis, personal fees from Regeneron/Sanofi, personal fees from Thermo Fisher Scientific, personal fees from Boehringer Ingelheim, personal fees from Novartis, personal fees from Philips, outside the submitted work. Dr. Belgrave reports personal fees from GSK, outside the submitted work.

Author contributions

JSK, EP, AS and CF contributed to the concept and design of the study. EP, EC, RF and AS contributed to the performance and analysis of preparation of samples and library for sequencing. EB, RF, EC and AS contributed to the performance and analysis of the qPCR data. SF and DB performed statistical analysis. All authors contributed to drafting the manuscript. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ebiom.2019.07.034.

References

[1] Bosch AA, Levin E, van Houten MA, et al. Development of upper respiratory tract microbiota in infancy is affected by mode of delivery. EBioMedicine 2016;9:336–45.
[2] Koenig JE, Spor A, Scalfone N, et al. Succession of microbial consortia in the developing infant gut microbiome. Proc Natl Acad Sci U S A 2011;108(Suppl. 1):4578–85.
[3] Abrahamsson TR, Jakobsson HE, Andersson AF, Bjorksten B, Engstrand L, Jenmalm MC. Low gut microbiota diversity in early infancy precedes asthma at school age. Clin Exp Allergy 2014;44(6):842–50.
[4] Russell SL, Gold MJ, Hartmann M, et al. Early life antibiotic-driven changes in microbiota enhance susceptibility to allergic asthma. EMBO Rep 2012;13(5):440–7.
[5] Hilty M, Burke C, Pedro H, et al. Disordered microbial communities in asthmatic airways. PLoS One 2010;5(1):e8578.
[6] Whiteson KL, Bailey B, Bergkessel M, et al. The upper respiratory tract as a microbial source for pulmonary infections in cystic fibrosis. Parallels from island biogeography. Am J Respir Crit Care Med 2014;189(11):1309–15.
[7] Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. The microbiome and the respiratory tract. Annu Rev Physiol 2016;78:481–504.
[8] Bisgaard H, Hermansen MN, Buchvald F, et al. Childhood asthma after bacterial colonization of the airway in neonates. New Engl J of Med 2007;357(15):1487–95.
[9] Teo SM, Mok D, Pham K, et al. The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development. Cell Host Microbe 2015;17(5):704–15.
[10] Dzidic M, Abrahamsson T, Artacho A, Collado MC, Mira A, Jenmalm MC. Oral microbiota maturation during the first 7 years of life in relation to allergy development. Allergy 2018;73(10):2000–11.
[11] Ta LDH, Yap GC, Tay CJX, et al. Establishment of the nasal microbiota in the first 18 months of life: correlation with early-onset rhinitis and wheezing. J Allergy Clin Immunol 2018;142(1):86–95.
[12] Depner M, Ege MJ, Cox MJ, et al. Bacterial microbiota of the upper respiratory tract and childhood asthma. J Allergy Clin Immunol 2017;139(3):826–34 e13.
[13] Boutin S, Depner M, Stahl M, et al. Comparison of Oropharyngeal microbiota from children with asthma and cystic fibrosis. Mediators Inflamm 2017;2017:5047403.
[14] Sim K, Cox MJ, Wopereis H, et al. Improved detection of bifidobacteria with optimised 16S rRNA-gene based pyrosequencing. PLoS One 2012;7(3):e32543.
[15] Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010;7(5):335–6.

[16] Pruesse E, Quast C, Knittel K, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 2007;35(21):7188–96.
[17] Yu Y, Lee C, Kim J, Hwang S. Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. Biotechnol Bioeng 2005;89(6):670–9.
[18] McMurdie PJ, Holmes S. Waste not, want not: why rarefying microbiome data is inadmissible. PLoS Comput Biol 2014;10(4):e1003531.
[19] McMurdie PJ, Holmes S. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 2013;8(4):e61217.
[20] FGB Jari Oksanen, Friendly Michael, Kindt Roeland, Legendre Pierre, PRM Dan McGlinn, O'Hara RB, et al. vegan: community ecology package. R package version 2.4-4; 2017.
[21] Paulson JN, Talukder H. Corrada Bravo H. longitudinal differential abundance analysis of microbial marker-gene surveys using smoothing splines. bioRxiv 2017. https://doi.org/10.1101/099457 099457.
[22] Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. Nat Methods 2013;10(12):1200–2.
[23] Paulson JN, Talukder H, Pop M, Bravo HC. metagenomeSeq: statistical analysis for sparse high-throughput sequencing. Bioconductor package: 1.19.0.
[24] de Steenhuijsen Piters WAA, Heinonen S, Hasrat R, et al. Nasopharyngeal microbiota, host Transcriptome, and disease severity in children with respiratory syncytial virus infection. Am J Respir Crit Care Med 2016;194(9):1104–15.
[25] Case RJ, Boucher Y, Dahllof I, Holmstrom C, Doolittle WF, Kjelleberg S. Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. Appl Environ Microbiol 2007;73(1):278–88.
[26] Deliu M, Belgrave D, Sperrin M, Buchan I, Custovic A. Asthma phenotypes in childhood. Expert Rev Clin Immunol 2016:1–9.
[27] Lowe L, Murray CS, Martin L, et al. Reported versus confirmed wheeze and lung function in early life. Arch Dis Child 2004;89(6):540–3.
[28] Pettigrew MM, Laufer AS, Gent JF, Kong Y, Fennie KP, Metlay JP. Upper respiratory tract microbial communities, acute otitis media pathogens, and antibiotic use in healthy and sick children. Appl Environ Microbiol 2012;78(17):6262–70.
[29] Hu Q, Dai W, Zhou Q, et al. Dynamic oropharyngeal and faecal microbiota during treatment in infants hospitalized for bronchiolitis compared with age-matched healthy subjects. Sci Rep 2017;7(1):11266.
[30] Stearns JC, Davidson CJ, McKeon S, et al. Culture and molecular-based profiles show shifts in bacterial communities of the upper respiratory tract that occur with age. ISME J 2015;9(5):1246–59.
[31] Bosch A, Piters W, van Houten MA, et al. Maturation of the infant respiratory microbiota, environmental drivers, and health consequences. A prospective cohort study. Am J Respir Crit Care Med 2017;196(12):1582–90.
[32] von Mutius E, Vercelli D. Farm living: effects on childhood asthma and allergy. Nat Rev Immunol 2010;10(12):861–8.
[33] Vissing NH, Chawes BL, Bisgaard H. Increased risk of pneumonia and bronchiolitis after bacterial colonization of the airways as neonates. Am J Respir Crit Care Med 2013;188(10):1246–52.
[34] Biesbroek G, Tsivtsivadze E, Sanders EA, et al. Early respiratory microbiota composition determines bacterial succession patterns and respiratory health in children. Am J Respir Crit Care Med 2014;190(11):1283–92.
[35] Cardenas PA, Cooper PJ, Cox MJ, et al. Upper airways microbiota in antibiotic-naive wheezing and healthy infants from the tropics of rural Ecuador. PLoS One 2012;7(10):e46803.
[36] Wang L, de Angel Sola D, Mao Y, et al. Family-based study reveals decreased abundance of sputum Granulicatella in asthmatics. Allergy 2018;73(9):1918–21.
[37] Rosas-Salazar C, Shilts MH, Tovchigrechko A, et al. Nasopharyngeal Lactobacillus is associated with a reduced risk of childhood wheezing illnesses following acute respiratory syncytial virus infection in infancy. J Allergy Clin Immunol 2018;142(5):1447–56.