Nasopharyngeal Lactobacillus is associated with a reduced risk of childhood wheezing illnesses following acute respiratory syncytial virus infection in infancy

Christian Rosas-Salazar, Vanderbilt University
Meghan H. Shiits, Vanderbilt University
Andrey Tovchigrechko, J. Craig Venter Institute
Seth Schobel, J. Craig Venter Institute
James D. Chappell, Vanderbilt University
Emma K. Larkin, Vanderbilt University
Tebeb Gebretsadik, Vanderbilt University
Rebecca A. Halpin, J. Craig Venter Institute
Karen E. Nelson, J. Craig Venter Institute
Martin Moore, Emory University

Only first 10 authors above; see publication for full author list.
Nasopharyngeal Lactobacillus is associated with a reduced risk of childhood wheezing illnesses following acute respiratory syncytial virus infection in infancy

Christian Rosas-Salazar, MD, MPH a,b,*, Meghan H. Shilts, MHS, MS c,e,g,*, Andrey Tovchigrechko, PhD d, Seth Schobel, PhD d, James D. Chappell, MD, PhD b, Emma K. Larkin, PhD b,c, Tebeb Gebretsadik, MPH b,d, Rebecca A. Halpin, MS a, Karen E. Nelson, PhD b, Martin L. Moore, PhD b,l, Larry J. Anderson, MD b,l, R. Stokes Peebles Jr, MD b,c, Suman R. Das, PhD c,e,g, and Tina V. Hartert, MD, MPH b,c

a Department of Pediatrics, Vanderbilt University School of Medicine, Nashville
b Center for Asthma Research, Vanderbilt University School of Medicine, Nashville
c Department of Medicine, Vanderbilt University School of Medicine, Nashville
d Department of Biostatistics, Vanderbilt University School of Medicine, Nashville
e Infectious Disease Group, J. Craig Venter Institute, Rockville
f Bioinformatics Group, J. Craig Venter Institute, Rockville
g Infectious Disease Group, J. Craig Venter Institute, La Jolla
h Genomic Medicine Group, J. Craig Venter Institute, La Jolla
i Department of Pediatrics, Emory University School of Medicine, Atlanta

Abstract

Background—Early life acute respiratory infection (ARI) with respiratory syncytial virus (RSV) has been strongly associated with the development of childhood wheezing illnesses, but the pathways underlying this association are poorly understood.

Objective—To examine the role of the nasopharyngeal microbiome in the development of childhood wheezing illnesses following RSV ARI in infancy.
Methods—We conducted a nested cohort study of 118 previously healthy, term infants with confirmed RSV ARI by RT-PCR. We used next-generation sequencing of the V4 region of the 16S ribosomal RNA gene to characterize the nasopharyngeal microbiome during RSV ARI. Our main outcome of interest was 2-year subsequent wheeze.

Results—Of the 118 infants, 113 (95.8%) had 2-year outcome data. Of these, 46 (40.7%) had parental report of subsequent wheeze. There was no association between the overall taxonomic composition, diversity, and richness of the nasopharyngeal microbiome during RSV ARI with the development of subsequent wheeze. However, the nasopharyngeal detection and abundance of Lactobacillus was consistently higher in infants who did not develop this outcome. Lactobacillus also ranked first among the different genera in a model distinguishing infants with and without subsequent wheeze.

Conclusions—The nasopharyngeal detection and increased abundance of Lactobacillus during RSV ARI in infancy are associated with a reduced risk of childhood wheezing illnesses at age 2 years.

Keywords
Microbiome; Lactobacillus; Staphylococcus; nasopharynx; respiratory syncytial virus; asthma; wheezing; 16S ribosomal RNA sequencing; infants

Respiratory syncytial virus (RSV) is one of the most common causes of upper and lower acute respiratory infections (ARIs) in young children worldwide. In the United States, it is associated with ~132,000 to ~172,000 inpatient admissions among preschool-aged children and its frequency appears to be increasing. In addition to its short-term effects, early life RSVARI, particularly severe infection in the critical period of infancy (ie, between 0 and 12 months of age), has been strongly associated with the development of childhood wheezing illnesses including asthma, the most common chronic disease of childhood. However, the pathways underlying this association are poorly understood.

In recent years, we and others have demonstrated that the upper airway bacterial microbiome plays an important role in the pathogenesis and pulmonary sequelae of viral ARIs. For example, several studies in murine models have shown that priming of the nasal mucosa with certain taxa (such as Lactobacillus species) increases the resistance and beneficially modulates the immune response of mice against RSV, influenza, and pneumonia virus. In a recent study of young children with RSV ARI, higher nasopharyngeal abundances of Haemophilus influenzae and Streptococcus were associated with a more proinflammatory immune response and a higher risk of hospitalization. In the same context, we have shown increased nasal abundances of Haemophilus, Moraxella, and Streptococcus and lower nasal abundances of Lactobacillus, Staphylococcus, and Corynebacterium in infants with RSV ARI when compared with healthy infants. Taken together, these findings suggest that viral-bacterial interactions have a crucial role in the severity of the ARI, the host immune response, and possibly the development of childhood wheezing illnesses.

On the basis of these findings, we hypothesized that the taxonomic composition, diversity, richness, and abundance of certain bacterial taxa of the upper airway microbiome during RSV ARI in infancy are associated with the later development of childhood wheezing.
illnesses. To test this hypothesis, we examined the association of the nasopharyngeal microbiome with childhood wheezing illnesses in infants with at least 1 RSV ARI enrolled in the Infant Susceptibility to Pulmonary Infections and Asthma Following RSV Exposure (INSPIRE) study.

METHODS

Overview of INSPIRE

INSPIRE is a current population-based birth cohort of previously healthy, term infants born between June and December of 2012 to 2013, designed so that the first RSV ARI during infancy could be studied. Eligible infants were enrolled mainly during a well-child visit at a participating general pediatric practice throughout the middle Tennessee region. The recruitment area encompasses urban, suburban, and rural areas. At enrollment, 1 of the parents was administered an extensive questionnaire to obtain information on the infant’s sociodemographic characteristics, birth and family history, and respiratory health. In order to capture an infant’s first RSV ARI, biweekly respiratory illness surveillance was performed during the winter viral season (November to March) of each infant’s first year of life. Infants who met prespecified criteria for an ARI had an in-person visit, which included a nasal wash for viral identification and characterization of the nasopharyngeal microbiome, as well as a physical examination for assessment of the ARI severity using the respiratory severity score (RSS). Annual follow-up to assess the development of childhood wheezing illnesses is ongoing. The Institutional Review Board of Vanderbilt University approved this study. The detailed methods for INSPIRE have been previously reported and are presented in detail in this article’s Online Repository at [www.jacionline.org](http://www.jacionline.org).

Study population

The current study included 118 infants with confirmed RSVARI enrolled in INSPIRE. A total of 125 nasal washes were included, as 5 infants had 2 RSV ARI episodes and 1 infant had 3. These 118 infants represent a nested cohort with nasopharyngeal microbiome assessment during a confirmed RSV ARI and available 2-year data (the most recent longitudinal follow-up visit available) at the time of the current study.

RSV detection

The detection of RSV was made by real-time RT-PCR. For this, TaqMan assays using RSV-specific primers and probes were run on the StepOnePlus platform (Applied Biosystems, Foster City, Calif) using the AgPath-ID OneStep RT-PCR Kit (Applied Biosystems) as per a previously described protocol.

Characterization of the nasopharyngeal microbiome

We have previously described in detail the methods used to characterize the nasopharyngeal microbiome using nasal washes in infants enrolled in INSPIRE. In brief, following bacterial DNA extraction, the V4 region of the 16S ribosomal RNA (rRNA) gene was amplified using universal 515F/806R primers. The libraries were then sequenced on an Illumina MiSeq platform with 2 × 300 bp reads. Negative and positive controls (with known
taxonomic composition) were amplified and sequenced concurrently for quality control. Further details on these steps are available in the Online Repository at www.jacionline.org.

**Outcome definitions and assessment**

Our outcome of interest was subsequent wheeze, defined as parental report of any wheeze since the last birthday. To test the robustness of our results, we also conducted sensitivity analyses using the outcome of recurrent wheeze, defined as parental report of ≥2 episodes of wheeze since the last birthday. Both of these outcomes were assessed at age 2 years (the most recent longitudinal follow-up visit available) using the International Study of Asthma and Allergy in Children questionnaire.22

**Data processing and statistical analyses**

A mothur-based automated annotation pipeline,23 YAP,24 was used to perform initial processing of the 16S rRNA gene sequencing datasets. Low-quality sequences, chimeras, and nonbacterial sequences are discarded as part of this pipeline. Samples with ≤1000 final reads (n = 1) were discarded prior to statistical analysis. Statistical analyses were performed with the open source MGSAT package in R.25,26 The MGSAT pipeline calls on a number of R tests and packages to compare the taxonomic composition, diversity, richness, and abundance of taxa between groups, including the permutation-based ANOVA test, Shannon index, inverse Simpson index, Chao1 estimator, observed taxa counts, DESeq2,27 stabsel,28,29 and GeneSelector.30 For analyses using operational taxonomic units (OTUs), OTUs were clustered at 97% sequence identity.

Our main method to test for differential abundance of taxa in association with childhood wheezing illnesses was DESeq2.27 DESeq2 models raw absolute counts of each taxon with a negative binomial distribution and uses the estimated depth of sequencing of each sample to scale the (unknown) relative abundance that is the parameter of the negative binomial distribution. Compared with using either simple proportion-based normalization or rarefaction for controlling for differential sequencing depth, the DESeq2 approach provides improved sensitivity and specificity.31 Reported Q values are the result of a Wald test with Benjamini and Hochberg correction for multiple comparisons.32 To build alternative rankings of taxa in regard to their importance in predicting the same childhood wheezing phenotype, we also used stabsel and GeneSelector. The stabsel stability selection approach aims to build the relative ranking of the predictor variables (taxa in our case) according to their importance for predicting the outcome.28 It does so by building multiple “base” models on random subsamples of the data (n = 400 in our study). We have used the elastic net model from the R package glmnet as the base feature selection method to be wrapped by the stability protocol.33 The ranking of taxa and their probability of being selected into the model were reported, as well as the probability cutoff corresponding to the per-family error rate that is controlled by this method. The GeneSelector package was used as a stability feature ranking method that is based on a nonparametric univariate test.30 In brief, the same ranking method (package function RankingWilcoxon) was applied to multiple random subsamples of the full set of observations (400 replicates, sampling 50% of observations without replacement). RankingWilcoxon ranks features in each replicate according to the test statistic from Wilcoxon rank-sum test with regard to the outcome group (eg, subsequent...
wheeze vs no subsequent wheeze). Consensus ranking between replicates was then found with a Monte Carlo procedure (package function AggregateMC) and the features were reported in the order of that consensus. To account for different sequencing depth, the absolute abundance counts were normalized to simple proportions within each observation. For each feature, we also obtained several types of the effect size, such as common language effect size and rank biserial correlation. If a taxon received similar ranking in all 3 statistical analyses (ie, DESeq2, stabsel, and GeneSelector), then findings for that particular taxon were considered robust and unlikely to occur due to chance.

For the DESeq2 analyses, we built both unadjusted and adjusted models. Because of their well-established association with the nasopharyngeal microbiome or childhood wheezing illnesses, our initial multivariable models include the following *a priori* selected covariates: infant’s age, sex, maternal asthma, and early life exposure to antibiotics (either *in utero* or after birth). Other models were then constructed by adding additional covariates, such as RSS or mode of delivery.

Based on our initial results, we also conducted several exploratory analyses to better understand the relation between early life nasopharyngeal colonization with *Lactobacillus* during RSV ARI and childhood wheezing illnesses. First, we examined the association of *Lactobacillus* detection (based on absolute counts) and relative abundance (simple proportions) with other infants’ baseline sociodemographic and clinical characteristics (ie, age, sex, race or ethnicity, gestational age, birth weight, mode of delivery, early life exposure to antibiotics, any breastfeeding, maternal smoking, maternal asthma, RSS, and type of insurance [a marker of socioeconomic status]). To this end, we used univariable logistic regression (for *Lactobacillus* detection) and proportional odds models (for *Lactobacillus* abundance) with robust sandwich standard error estimation (Huber-White method) to account for subjects’ cluster. Next, to obtain an estimate of the magnitude of the association and further evaluate the possibility of confounding, we used multivariable logistic regression to assess the association between *Lactobacillus* detection with the subsequent and recurrent wheeze outcome definitions while adjusting for age, sex, maternal asthma, and early-life exposure for antibiotics. Other models were then constructed by adding additional covariates, such as RSS or mode of delivery. The Huber-White method was also used to account for subjects’ cluster.

Statistical significance was defined as $P < .05$ after controlling for multiple comparisons when appropriate. Further details on the statistical analyses are available in the Online Repository at [www.jacionline.org](http://www.jacionline.org).

**RESULTS**

**Baseline characteristics of the study population**

The baseline characteristics of the 118 infants with RSV ARI included in this study (as a whole and according to the outcomes of interest) are presented in Table I and in Table E1 in this article’s Online Repository at [www.jacionline.org](http://www.jacionline.org). The median age at the time of the RSV ARI was 21.8 (interquartile range [IQR]: 12.1–27.1) weeks. A total of 113 (95.8%) of the 118 infants had 2-year outcome data. Of these, 46 (40.7%) and 36 (31.9%) had parental...
report of subsequent and recurrent wheeze, respectively. Infants with subsequent and recurrent wheeze were more likely to be male, to be born by cesarean section, to have been exposed to antibiotics in early life, to have a nonsmoking mother, and to have a maternal history of asthma when compared with those without these outcomes, although none of these differences was statistically significant \( (P > .05) \) for all estimates. Infants with recurrent wheeze had a significantly higher RSS when compared with those without \( (P = .03) \).

**Overall characteristics of the nasopharyngeal microbiome**

The total high-quality sequence count was 2,138,976 with a median sequence count per sample of 18,130 (IQR: 13,240–25,970). These sequences represented a total of 357 bacterial genera, with a median of 19 (IQR: 13–31) estimated observed genera per sample. The overall taxonomic composition of the nasopharyngeal microbiome during RSV ARI in infancy was characterized by high relative abundances of *Moraxella* (37.6%), *Streptococcus* (19.7%), *Haemophilus* (13.5%), *Corynebacterium* (10.0%), and *Dolosigranulum* (4.7%), and low relative abundances of the remaining bacterial genera (with a combined relative abundance of 14.5%).

**Main analyses**

**The overall taxonomic composition, diversity, and richness of the nasopharyngeal microbiome during RSV ARI in infancy were not associated with the development of subsequent wheeze**—The overall taxonomic composition of the nasopharyngeal microbiome between infants with and without subsequent wheeze did not differ at the OTU \( (P = .9) \) or genus level \( (P = .7) \) in the permutation-based ANOVA test of pairwise Bray-Curtis dissimilarities. There was a trend toward higher diversity and richness in infants with subsequent wheeze using the Shannon index, inverse Simpson index, Chao1 estimator, and observed taxa counts at both the OTU (Fig 1, A) and genus (Fig 1, B) levels, although these did not reach statistical significance \( (P > .05) \) for all estimates.

**Nasopharyngeal *Lactobacillus* was associated with reduced risk of subsequent wheeze following RSV ARI in infancy**—In the initial DESeq2 analyses, the absolute counts of 3 genera were detected to be significantly different according to the development of subsequent wheeze: the absolute counts of *Lactobacillus* (DESeq2 test base mean = 4.93; log\(_2\) fold change = −5.21; \( Q = 4.69\times10^{-8} \)) and *Staphylococcus* (DESeq2 test base mean = 59.29; log\(_2\) fold change = −1.47; \( Q = 2.08\times10^{-2} \)) were lower in infants with subsequent wheeze compared with those without, whereas the absolute counts of *Pseudomonas* (DE-Seq2 test base mean = 5.84; log\(_2\) fold change = 2.22; \( Q = 4.95\times10^{-4} \)) were higher. To control for potential confounders of these associations, we built a set of alternative DESeq2 models while adjusting for different relevant covariates as described in the Methods section. Only the changes in *Lactobacillus* and *Staphylococcus* absolute counts remained statistically significant in all models (Table II).

The relative abundances of the 35 most common genera expressed as simple proportions and split according to the development of subsequent wheeze are shown in Fig 2. *Lactobacillus* was the 31st most abundant genus, while *Staphylococcus* was the 15th. The mean ± SD
relative abundances of Lactobacillus and Staphylococcus over all samples where subsequent wheeze data was available were 9.39e-04 ± 6.53e-03 and 0.02 ± 0.11, respectively. In spite of its low relative abundance, the mean ± SD relative abundance of Lactobacillus was ~2 orders of magnitude lower in infants with subsequent wheeze compared with those without (2.47e-05 ± 5.51e-05 vs 1.94e-03 ± 1.94e-03; P = .04 using a Mann-Whitney U-test). In contrast, there was no significant difference in the mean ± SD relative abundance of Staphylococcus between infants with and without subsequent wheeze (6.89e-03 ± 1.84e-02 vs 3.67e-02 ± 1.52e-01; P = .9 using a Mann-Whitney U-test).

In a stabsel stability selection protocol distinguishing infants with and without subsequent wheeze using elastic net regression, Lactobacillus ranked first among the different genera, with a probability of being included into the model of 0.97, while Staphylococcus ranked second, with a probability of being included into the model of 0.76 (Fig 3). The mean relative abundance and standard deviation of the top 20 ranked genera in the stability selection model according to subsequent wheeze are shown in Table E2 in this article’s Online Repository at www.jacionline.org.

In a GeneSelector stability ranking procedure that wraps a nonparametric Wilcoxon rank-sum test, Lactobacillus was again ranked first among all genera (rank-biserial correlation effect size of subsequent wheeze relative to the group without subsequent wheeze = −0.23), while Staphylococcus was ranked 32nd (rank-biserial correlation effect size of subsequent wheeze relative to the group without subsequent wheeze = 0.01).

To further evaluate the possibility of confounding by RSV ARI severity, we then examined the abundance of Lactobacillus and Staphylococcus in each subsequent wheeze group over the different values of the RSS. The relative abundance of Lactobacillus was generally lower among infants with subsequent wheeze over all values of the RSS (Fig 4). In contrast, there was no longer a clear trend for Staphylococcus, which suggests a potential confounding effect of RSV ARI severity for this particular genus (see Fig E1 in this article’s Online Repository at www.jacionline.org).

We performed several additional analyses to examine the reliability of the association for Lactobacillus and Staphylococcus with subsequent wheeze, which can be found in the Online Repository at www.jacionline.org. In these analyses, none of the infants with a relative abundance of Lactobacillus greater than ~0.001 developed subsequent wheeze when using either unrarefied or rarified datasets (see Figs E2 and E3 in this article’s Online Repository at www.jacionline.org).

Sensitivity analyses

In spite of the smaller sample size of infants with recurrent wheeze, we found similar results when using this outcome instead of subsequent wheeze. The overall taxonomic composition, diversity, and richness of the nasopharyngeal microbiome did not differ at the genus level in infants with RSVARI according to the development of recurrent wheeze (P > .05 for all estimates); however, infants with recurrent wheeze also had significantly lower Lactobacillus abundance than did those without in DESeq2 analyses (DESeq2 test base mean = 2.81; log2 fold change = −3.58; Q = 1.47e-03), and the relative abundance of Lactobacillus was also

J Allergy Clin Immunol. Author manuscript; available in PMC 2018 November 09.
generally lower among infants with recurrent wheeze over all values of the RSS (see Fig E4 in this article’s Online Repository at [www.jacionline.org](http://www.jacionline.org)).

**Exploratory analyses**

In exploratory analyses, there was no association between *Lactobacillus* detection or abundance and any of the covariates shown in Table I (*P* > .05 for all estimates). The genus *Lactobacillus* was detected in 39 of the 125 samples (31.2%) and the proportion of samples with this genus was lower in infants with subsequent wheeze than in those without (26% vs 49%). In a multivariable model adjusting for age, sex, maternal asthma, and early life exposure for antibiotics, the detection of *Lactobacillus* decreased the odds of subsequent wheeze by ~70% (odds ratio [OR]: 0.34; 95% CI:0.14–0.83; *P* = .02) and of recurrent wheeze by ~80% (OR: 0.21; 95% CI: 0.07–0.65; *P* = .006). Including other potential confounders in these multivariable models, such as RSS or mode of delivery, or conducting the exploratory analyses after rarefication to the lowest library size found among all samples with the nonzero initial count of *Lactobacillus* did not significantly change our results (data not shown).

**DISCUSSION**

In our study, we found that the nasopharyngeal detection and increased abundance of *Lactobacillus* during RSV ARI in infancy are associated with a reduced risk childhood wheezing illnesses at age 2 years. Despite its low abundance compared with other taxa, *Lactobacillus* was still identified as being the strongest, most consistent discriminating taxon between infants with and without childhood wheezing illnesses by age 2 years across several different statistical methods. Furthermore, our results suggest that the detection of *Lactobacillus* in the nasopharynx of RSV-infected infants could be used as a biomarker for the later development of childhood wheezing illnesses. There was also a trend toward higher diversity and richness of the nasopharyngeal microbiome during RSV ARI in infants who developed childhood wheezing illnesses at age 2 years, although this was not statistically significant. We have previously shown a higher nasopharyngeal microbial diversity and richness in infants born via cesarean section compared with infants born via vaginal delivery. A higher diversity and richness of the gut microbiome has been associated with better outcomes, whereas a higher microbial diversity and richness of the vaginal microbiome has been associated with worse outcomes. Thus, the relationship of diversity and richness with health appears to be specific to each human body habitat.

One other study has examined the association of the early life respiratory microbiome and childhood wheezing illnesses using culture-independent, next-generation sequencing techniques. In this study, the investigators found that a high nasopharyngeal abundance of *Streptococcus* in infancy was a strong predictor of current wheeze at age 5 years but not at age 10 years. Unlike ours, this study assessed the upper airway microbiome during health and included only children born to atopic parents. In addition, the nasopharyngeal detection and abundance of *Lactobacillus* were not reported and the statistical analyses used focused only on the most abundant genera, which could explain the contradictory results.
To our knowledge, this is the first study to show a protective effect of upper airway colonization with *Lactobacillus* on asthma-related outcomes in children. In biologic support of our findings, Tomosada et al\(^9\) previously showed that nasal administration of *Lactobacillus rhamnosus* improves the antiviral immune response, enhances viral clearance, and reduces the lung injury associated with RSV ARI in mice. In humans, 1 study found a low abundance of *Lactobacillus sakei* in the maxillary sinus in adults with chronic rhinosinusitis,\(^38\) while another found a low abundance of the genus *Lactobacillus* in the trachea of premature newborns who later developed bronchopulmonary dysplasia.\(^39\) In the same context, we have previously shown a higher nasopharyngeal abundance of *Lactobacillus* in healthy infants when compared with infants with RSV ARI.\(^13\) *Lactobacilli* are also the primary bacteria of the vaginal flora of healthy women, and birth by cesarean section (which likely reduces upper airway colonization with *Lactobacillus* species) has been repeatedly associated with an increased risk of childhood wheezing illnesses.\(^35\),\(^40\)–\(^42\) Taken together, these findings strongly suggest a protective role of upper airway colonization with *Lactobacillus* in childhood respiratory outcomes.

While the effect of upper airway colonization with *Lactobacillus* on asthma-related outcomes in both children and adults has been understudied, numerous studies have focused on the effect of gastrointestinal supplementation with *Lactobacillus* species on childhood wheezing illnesses, yielding conflicting results.\(^43\) Because the nasopharyngeal route can induce local and systemic immune responses superior to those obtained using the oral route, \(^9,\)^10 respiratory supplementation with *Lactobacillus* in early life could be a potential strategy to prevent childhood wheezing illnesses.

In addition to *Lactobacillus*, we also found an association of upper airway colonization with *Staphylococcus* on childhood wheezing illnesses following RSV ARI in infancy. However, unlike *Lactobacillus*, this effect was not consistent across all statistical analyses and this lack of consistency (described in greater detail in the Online Repository at [www.jacionline.org](http://www.jacionline.org)) prevents us from confidently concluding that *Staphylococcus* is protective against childhood wheezing illnesses within this cohort. *Staphylococcus* is one of the most common colonizers in the nasopharynx of healthy infants and its abundance appears to decrease over time.\(^14,\)^16,\(^44,\)^45 We have previously shown a higher nasopharyngeal abundance of *Staphylococcus* in healthy infants when compared with infants with RSV ARI.\(^13\) In the only other study examining the association of the early life respiratory microbiome and childhood wheezing illnesses using next-generation sequencing, no association was noted between nasopharyngeal colonization with *Staphylococcus* in infancy and current wheeze at age 5 or 10 years.\(^14\)

The gut microbiome has been shown to play an important role in preserving host health.\(^46,\)^47 More recently, the manipulation of the gut microbiome has emerged as a strategy to prevent and treat certain diseases.\(^48,\)^49 In contrast to the gut microbiome, the upper airway microbiome has been relatively understudied. However, the upper airway is of germane interest in viral ARIs and childhood allergic diseases, such as asthma, as it is the first portal of entry of respiratory pathogens and aeroallergens. Our findings of nasopharyngeal bacterial genera independently associated with childhood wheezing illnesses are meaningful, as the implications are that targeted modifications of the infant upper airway microbiome
could be a strategy to decrease morbidity from both RSV ARIs and childhood asthma, the most common acute and chronic diseases of infancy and childhood, respectively.

Our study has considerable strengths, including the population-based longitudinal design of the parent study, the inclusion of infants with different degrees of RSV ARI severity, the use of predefined criteria for ARIs, and the close surveillance during the first winter viral season to capture each infant’s initial RSV ARI. We also acknowledge several limitations. First, due to the inherent limits of 16S rRNA gene sequencing-based analyses, we were unable to identify bacterial taxa below the genus level and thus cannot identify whether particular Lactobacillus strains or species, or a consortium of members of the Lactobacillus genus, protect against childhood wheezing illnesses. The V4 region of the 16S rRNA gene is not sufficient to identify Lactobacillus species and our OTU-based analyses were also limited due to the low abundance of Lactobacillus in the nasopharynx of infants with RSV ARI. Unlike the genus Lactobacillus, which contains species generally regarded as safe and beneficial for human health, the genus Staphylococcus contains species that can be potentially detrimental and, thus, translation of our results into preventive strategies for this particular genus should be done with caution. Second, the nasopharyngeal abundance of Lactobacillus was generally low, which may have affected our estimates. Third, due to the observational design, there could also be residual confounding by variables not measured in our study (such as prior infection or coinfection with other respiratory viruses). Fourth, because childhood wheezing illnesses are a heterogeneous group of disorders,50 the effect of upper airway colonization with Lactobacillus on later asthma phenotypes will need to be examined. Last, our results might not be generalizable to infants without RSV ARI. However, RSV is a ubiquitous infection and a major cause of infant morbidity worldwide, thus our findings are relevant to a group at a high risk of developing childhood wheezing illnesses.

In summary, we found that the nasopharyngeal detection and increased abundance of Lactobacillus during RSV ARI in infancy are associated with a reduced risk of childhood wheezing illnesses at age 2 years. While these preliminary findings merit replication in larger longitudinal cohorts, they provide novel data that could support the development of new prognostic or preventive strategies for childhood wheezing illnesses.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Acknowledgments**

This work was supported in whole or in part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under award numbers U19AI095227, K24AI77930, HHSN272200900007C, and U19AI110819; the Vanderbilt Institute for Clinical and Translational Research grant support (National Center for Advancing Translational Sciences, National Institutes of Health, Department of Health and Human Services, under award numbers UL1 TR000445 and U54RR24975); the Vanderbilt Faculty Research Scholars Program; and the Parker B. Francis Fellowship Program.

**Abbreviations used**

| ARI | Acute respiratory infection |

*J Allergy Clin Immunol. Author manuscript; available in PMC 2018 November 09.*
INSPIRE Infant Susceptibility to Pulmonary Infections and Asthma Following RSV Exposure

IQR Interquartile range

OTU Operational taxonomic unit

rRNA Ribosomal RNA

RSS Respiratory severity score

RSV Respiratory syncytial virus

References


### Key messages

- The nasopharyngeal detection and increased abundance of *Lactobacillus* during ARI with RSV in infancy are associated with a reduced risk of childhood wheezing illnesses at age 2 years.

- Our results suggest that the detection of *Lactobacillus* in the nasopharynx of RSV-infected infants could be used as a biomarker for the later development of childhood wheezing illnesses.

- Furthermore, providing *Lactobacillus* during RSV ARI in infancy could be a potential primary prevention intervention strategy for the development of childhood wheezing illnesses.
FIG. 1.
Boxplots of nasopharyngeal diversity and richness in infants with RSV ARI at the OTU (A) or genus (B) level by 2-year subsequent wheeze. After rarefaction to the lowest library size, \( \alpha \) diversity and richness estimates were calculated per each sample. This process was repeated 400 times and results were averaged. The Shannon and inverse Simpson indices were calculated to estimate abundance-based OTU or genus diversity, while the Chao1 estimator and observed taxa counts were calculated to estimate abundance-based OTU or genus richness. Both OTU and genus richness and diversity were lower in infants without 2-year subsequent wheeze, although this was not significant in any case.
FIG. 2.
Boxplots of relative abundance of nasopharyngeal bacterial genera in infants with RSV ARI by 2-year subsequent wheeze. Within each sample, counts were normalized to simple proportions. The relative abundance of the 35 most abundant genera is shown; all other genera are not shown in this figure. The median (middle bar), third quartile (right-most bar), and first quartile (left-most bar) abundances are shown. Outliers are represented as dots.
FIG. 3.
Probability of a nasopharyngeal bacterial genus being selected into a stability selection model distinguishing infants with RSV ARI with and without 2-year subsequent wheeze. The probability is plotted along the x-axis. The genera are indicated along the y-axis. The top 20 ranked genera are shown in this figure. Lactobacillus ranks highest among taxa selected into the model with a probability of being selected of 0.97.
FIG. 4.
Box-Cox transformed nasopharyngeal relative abundance of *Lactobacillus* in infants with RSV ARI with *(blue line)* and without *(red line)* 2-year subsequent wheeze, plotted by RSS. Lines are local regression (LOESS) smoothed curves and gray areas are the 95% CIs. For the y-axis, values closer to 0 indicate a higher abundance. Not all individual data points are shown; a single data point is displayed for infants who had the same RSS and *Lactobacillus* abundance. In general, the relative abundance of *Lactobacillus* was lower among infants with subsequent wheeze over all values of the RSS.
### TABLE I

Baseline characteristics of infants with RSV ARI included in this study (n = 118)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (wk)</td>
<td>21.8 (12.1–27.1)</td>
</tr>
<tr>
<td>Female sex</td>
<td>50 (42.4)</td>
</tr>
<tr>
<td>Race or ethnicity</td>
<td></td>
</tr>
<tr>
<td>Black non-Hispanic</td>
<td>20 (17.0)</td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>74 (62.7)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>11 (9.3)</td>
</tr>
<tr>
<td>Other *</td>
<td>13 (11.0)</td>
</tr>
<tr>
<td>Gestational age (wk)</td>
<td>39.0 (38.5–40.0)</td>
</tr>
<tr>
<td>Birth weight (g)</td>
<td>3377 (2894–3859)</td>
</tr>
<tr>
<td>Birth by cesarean section</td>
<td>42 (35.6)</td>
</tr>
<tr>
<td>Exposure to antibiotics in utero or after birth</td>
<td>62 (52.5)</td>
</tr>
<tr>
<td>Any breastfeeding</td>
<td>86 (72.9)</td>
</tr>
<tr>
<td>Maternal smoking at enrollment</td>
<td>26 (22.0)</td>
</tr>
<tr>
<td>Maternal asthma</td>
<td>20 (17.0)</td>
</tr>
<tr>
<td>Respiratory severity score</td>
<td>3.0 (2.0–5.0)</td>
</tr>
<tr>
<td>Insurance type</td>
<td></td>
</tr>
<tr>
<td>Medicaid</td>
<td>59 (50.0)</td>
</tr>
<tr>
<td>Private</td>
<td>56 (47.5)</td>
</tr>
<tr>
<td>Other</td>
<td>3 (2.5)</td>
</tr>
</tbody>
</table>

Data are presented as median (interquartile range) for continuous variables or n (%) for binary variables. Percentages calculated for children with complete data.

* Other includes mixed race and unknown.
<table>
<thead>
<tr>
<th>Genus</th>
<th>Model 1*</th>
<th></th>
<th></th>
<th></th>
<th>Model 2*</th>
<th></th>
<th></th>
<th></th>
<th>Model 3*</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base mean</td>
<td>Log₂-fold change (SE)**</td>
<td>Q value</td>
<td>Base mean</td>
<td>Log₂-fold change (SE)**</td>
<td>Q value</td>
<td>Base mean</td>
<td>Log₂-fold change (SE)**</td>
<td>Q value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lactobacillus</td>
<td>8.24</td>
<td>-4.45 (0.73)</td>
<td>3.25e-08</td>
<td>8.42</td>
<td>-4.56 (0.74)</td>
<td>3.22e-08</td>
<td>8.24</td>
<td>-4.57 (0.74)</td>
<td>2.02e-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staphylococcus</td>
<td>901.28</td>
<td>-2.59 (0.57)</td>
<td>1.00e-04</td>
<td>947.73</td>
<td>-2.94 (0.58)</td>
<td>6.75e-06</td>
<td>901.28</td>
<td>-2.84 (0.57)</td>
<td>1.36e-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudomonas</td>
<td>53.98</td>
<td>2.29 (0.60)</td>
<td>1.39e-03</td>
<td>56.34</td>
<td>2.39 (0.61)</td>
<td>9.17e-04</td>
<td>†</td>
<td>†</td>
<td>†</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data presented are the results of the DESeq2 test. Base means are calculated after normalizing read counts for each sample to account for differences in sequencing depth. Reported Q values are the result of a Wald test with Benjamini and Hochberg correction for multiple comparisons. Genera are ordered from lowest to highest Q value. Only genera with Q values below the statistical significance threshold of .05 in the initial DESeq2 test are presented (see main text for details).

* Model 1 includes age, sex, maternal asthma, and exposure to antibiotics in utero or after birth as covariates. Model 2 includes the same covariates as model 1 plus respiratory severity score. Model 3 includes the same covariates as model 1 plus mode of delivery.

** A log₂-fold change of >0 indicates that the abundance of that particular genus was detected to be higher in infants with RSV ARI who later developed subsequent wheeze when compared with those who did not develop this outcome, while a log₂-fold change <0 indicates the opposite.

† Indicates Q value was not <.05 for the genus in the specified model.