

**Age-related changes in the nasopharyngeal microbiome are associated with SARS-CoV-2 infection
and symptoms among children, adolescents, and young adults**

Jillian H. Hurst, PhD^{1,2,†}, Alexander W. McCumber, PhD^{1,†}, Jhoanna N. Aquino, BS¹, Javier Rodriguez, AS³, Sarah M. Heston, MD¹, Debra J. Lugo, MD¹, Alexandre T. Rotta, MD⁴, Nicholas A. Turner, MD, MHSc⁵, Trevor S. Pfeiffer, BS¹, Thaddeus C. Gurley, MS⁶, M. Anthony Moody, MD^{1,6}, Thomas N. Denny, MSc, MPhil⁶, John F. Rawls, PhD^{7,8}, James S. Clark, PhD⁹, Christopher W. Woods, MD, MPH^{5,6}, Matthew S. Kelly, MD, MPH^{1,8*}

¹Department of Pediatrics, Division of Infectious Diseases, Duke University School of Medicine, Durham, North Carolina, USA; ²Children's Health and Discovery Institute, Department of Pediatrics, Duke University School of Medicine, Durham, North Carolina, USA; ³Children's Clinical Research Unit, Department of Pediatrics, Duke University School of Medicine, Durham, North Carolina, USA; ⁴Department of Pediatrics, Division of Pediatric Critical Care Medicine, Duke University School of Medicine, Durham, North Carolina, USA; ⁵Department of Medicine, Division of Infectious Diseases, Duke University School of Medicine, Durham, North Carolina, USA; ⁶Duke Human Vaccine Institute, Duke University School of Medicine, Durham, North Carolina, USA; ⁷Department of Molecular Genetics and Microbiology, Duke University School of Medicine, Durham, North Carolina, USA; ⁸Duke Microbiome Center, Duke University School of Medicine, Durham, North Carolina, USA; ⁹Nicholas School of the Environment, Duke University, Durham, North Carolina, USA

[†]these authors contributed equally to this work; *corresponding author

© The Author(s) 2022. Published by Oxford University Press for the Infectious Diseases Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Corresponding Author:

Matthew S. Kelly, MD, MPH

2301 Erwin Road

Durham, NC 27710 USA

Email: matthew.kelly@duke.edu

Summary: We demonstrate that the nasopharyngeal microbiome undergoes marked shifts in composition with age during childhood and adolescence, and that age-associated changes in nasopharyngeal microbiome composition are associated with SARS-CoV-2 infection and SARS-CoV-2-associated respiratory symptoms among children, adolescents, and young adults.

Accepted Manuscript

ABSTRACT

Background: Children are less susceptible to SARS-CoV-2 infection and typically have milder illness courses than adults, but the factors underlying these age-associated differences are not well understood. The upper respiratory microbiome undergoes substantial shifts during childhood and is increasingly recognized to influence host defense against respiratory pathogens. Thus, we sought to identify upper respiratory microbiome features associated with SARS-CoV-2 infection susceptibility and illness severity.

Methods: We collected clinical data and nasopharyngeal swabs from 285 children, adolescents, and young adults (<21 years of age) with documented SARS-CoV-2 exposure. We used 16S ribosomal RNA gene sequencing to characterize the nasopharyngeal microbiome and evaluated for age-adjusted associations between microbiome characteristics and SARS-CoV-2 infection status and respiratory symptoms.

Results: Nasopharyngeal microbiome composition varied with age (PERMANOVA, $p < 0.001$, $R^2 = 0.06$) and between SARS-CoV-2-infected individuals *with* and *without* respiratory symptoms (PERMANOVA, $p = 0.002$, $R^2 = 0.009$). SARS-CoV-2-infected participants with *Corynebacterium/Dolosigranulum*-dominant microbiome profiles were less likely to have respiratory symptoms than infected participants with other nasopharyngeal microbiome profiles (odds ratio: 0.38, 95% confidence interval: 0.18–0.81). Using generalized joint attributed modeling, we identified nine bacterial taxa associated with SARS-CoV-2 infection and six taxa that were differentially abundant among SARS-CoV-2-infected participants with respiratory symptoms; the magnitude of these associations was strongly influenced by age.

Conclusions: We identified interactive relationships between age and specific nasopharyngeal microbiome features that are associated with SARS-CoV-2 infection susceptibility and symptoms in children, adolescents, and young adults. Our data suggest that the upper respiratory microbiome may be a mechanism by which age influences SARS-CoV-2 susceptibility and illness severity.

Keywords: COVID-19; pediatric microbiota; *Corynebacterium*; *Dolosigranulum*; generalized joint attribute modeling

Accepted Manuscript

INTRODUCTION

In contrast to most other respiratory viruses [1], children appear to be less susceptible to infection with severe acute respiratory virus 2 (SARS-CoV-2), and typically have milder illness courses than adults. In a recent meta-analysis of 32 studies that included 41,640 children and adolescents and 268,945 adults, SARS-CoV-2 infection susceptibility was estimated to be 46% lower among children and adolescents relative to adults [2]. Further, a higher incidence of SARS-CoV-2 infection has been observed with increasing age, even among infants, children, and adolescents [3]. We previously demonstrated that up to one-third of SARS-CoV-2-infected children and adolescents are asymptomatic [4], and the vast majority of children who develop symptoms report mild respiratory symptoms [4, 5]. Additionally, COVID-19 hospitalization rates and mortality among children are substantially lower than among adults of all ages [6]. These data suggest that changes in host biological or immunological factors that occur with age modify susceptibility to and severity of SARS-CoV-2 infection.

Given that the upper respiratory microbiome undergoes substantial shifts in early childhood [7, 8], and is increasingly recognized to play a key role in the pathogenesis of respiratory virus infections [9, 10], we hypothesized that age-associated changes in the upper respiratory microbiome might contribute to differences in SARS-CoV-2 susceptibility and illness severity among children and adults. In this study, we used 16S ribosomal RNA (rRNA) gene amplicon sequencing to characterize the nasopharyngeal microbiomes of 285 children, adolescents, and young adults with close contact with a SARS-CoV-2-infected individual and to identify microbiome features associated with SARS-CoV-2 infection and with the presence of respiratory symptoms among SARS-CoV-2-infected individuals.

METHODS

Study procedures

The Duke Biospecimens from Respiratory Virus-Exposed Kids (BRAVE Kids) study is a prospective cohort study of children, adolescents, and young adults (<21 years of age) with confirmed SARS-CoV-2 infection or close contact with an individual with confirmed SARS-CoV-2 infection, as previously described [4]. Exposure, sociodemographic, and clinical data are collected at enrollment, and we record symptoms occurring up to 14 days prior to and 28 days after study enrollment. Nasopharyngeal samples are collected with nylon flocked swabs (Copan Italia, Brescia, Italy) and placed into RNAProtect (Qiagen, Hilden, Germany) prior to storage at -80°C. Participants are classified as SARS-CoV-2-infected if the virus is detected in either a clinical or research PCR assay. For the analyses presented herein, we considered SARS-CoV-2-infected individuals to have respiratory symptoms if they reported cough, rhinorrhea, nasal congestion, shortness of breath, sore throat, or anosmia at any point between 14 days prior to enrollment through 28 days after enrollment.

Processing of nasopharyngeal samples for 16S ribosomal RNA sequencing

The Duke Microbiome Core Facility extracted DNA from nasopharyngeal samples using Powersoil Pro extraction kits (Qiagen). DNA concentrations were determined using Qubit dsDNA high-sensitivity assay kits (Thermo Fisher Scientific). Bacterial community composition was characterized by PCR amplification of the V4 variable region of the 16S rRNA gene [11]. Equimolar 16S rRNA PCR products were quantified and pooled prior to sequencing. Sequencing was performed by the Duke Sequencing and Genomic Technologies Core Facility on an Illumina MiSeq instrument configured for 250 base-pair paired-end sequencing. All samples were included in a single sample processing run with negative extraction and PCR controls. We analyzed raw sequences using DADA2 version 1.16 [12]

and assigned taxonomy to amplicon sequence variants (ASVs) using the expanded Human Oral Microbiome Database version 15.1 [13]. We identified and removed presumed reagent contaminant ASVs (n=35; **Supplemental Table 1**) based on presence in negative control samples or negative correlation with DNA concentration using the frequency method (threshold=0.10) implemented in the *decontam* R package version 1.12 [14]. We excluded samples with less than 1,000 sequencing reads after quality filtering and contaminant removal. We obtained a median [interquartile range (IQR)] of 24,360 (18,117–33,371) high-quality sequencing reads from the 285 samples included in these analyses. Sequencing reads were classified into 1,854 ASVs representing 202 bacterial genera from 8 phyla.

Data analysis

We calculated nasopharyngeal microbiome alpha diversity (Shannon diversity index and number of unique ASVs) using the *phyloseq* R package version 1.36 [15]. We fit linear regression models to evaluate associations between patient characteristics and microbiome alpha diversity measures. The number of unique ASVs was not normally distributed and was log-transformed for these analyses. We used the *microbiome* R package version 1.8.0 [16] to generate centered log-ratio (CLR)-transformed sample counts to evaluate between-sample compositional differences [17]. We used k-medoids clustering and the Calinski-Harabasz index to classify samples into distinct nasopharyngeal microbiome profiles. We evaluated associations between patient characteristics and nasopharyngeal microbiome composition with PERMANOVA using the *adonis* function within the *vegan* R package version 2.5.7 [18]. To evaluate associations between patient characteristics and the relative abundances of specific ASVs within the nasopharyngeal microbiome, we used generalized joint attribute modeling (GJAM) implemented in the *gjam* R package version 2.3.5 [19]. Analyses conducted in *gjam* were limited to ASVs present in at least 5% of samples. We adjusted for participant age (as a continuous variable) and assessed the significance of interaction terms in all

analyses to evaluate for interactive relationships between age and the relative abundances of specific ASVs on SARS-CoV-2 infection and SARS-CoV-2-associated respiratory symptoms. Our findings in all analyses were not substantively changed when we additionally adjusted for sex and race (data not shown). All analyses were conducted in R version 4.1 [20].

RESULTS

Patient characteristics

Two hundred eighty-five children, adolescents, and young adults were included in these analyses (**Table 1**). Participants were classified as SARS-CoV-2-exposed but uninfected (n=74, 26%); SARS-CoV-2-infected *without* respiratory symptoms (n=98, 34%); and SARS-CoV-2-infected *with* respiratory symptoms (n=113, 40%). SARS-CoV-2-infected participants *with* respiratory symptoms were older than SARS-CoV-2-infected participants *without* respiratory symptoms [median (IQR) age: 14.1 (6.3–17.5) vs. 9.3 (4.8–13.2) years; Wilcoxon rank-sum test, $p=0.001$] and SARS-CoV-2-uninfected participants [median (IQR) age: 14.1 (6.3–17.5) vs. 9.5 (5.1–15.8) years; Wilcoxon rank-sum test, $p=0.051$]. There were no significant differences in the prevalences of comorbidities or recent receipt of antibiotics or probiotics in these groups.

Nasopharyngeal microbiome diversity differs based on age and SARS-CoV-2 infection status

We first sought to describe changes in nasopharyngeal microbiome diversity that occur with age from infancy through early adulthood, and by SARS-CoV-2 infection and symptom status. Median (IQR) Shannon diversity and number of unique ASVs in nasopharyngeal samples were 1.49 (1.12–2.00) and 67 (48–90), respectively. Nasopharyngeal microbiome diversity, as measured by the

Shannon index, increased with increasing age (**Figure 1A**; ANCOVA, $p < 0.0001$), while observed richness (number of unique ASVs) was not associated with participant age (**Figure 1B**; ANCOVA, $p = 0.27$). Shannon diversity was similar in nasopharyngeal samples from SARS-CoV-2-infected and uninfected participants [median (IQR): 1.49 (1.16–2.02) vs. 1.64 (1.20–2.04); ANCOVA, $p = 0.39$]; however, observed richness was higher in SARS-CoV-2-infected participants than in SARS-CoV-2-uninfected participants [median (IQR): 69 (51–75) vs. 59 (38–84); ANCOVA, $p = 0.01$]. Measures of nasopharyngeal microbiome diversity were similar in SARS-CoV-2-infected individuals *with* or *without* respiratory symptoms (data not shown).

Nasopharyngeal microbiome composition by age, SARS-CoV-2 infection, and COVID-19 symptoms

Five bacterial genera accounted for more than 80% of the sequencing reads identified in nasopharyngeal samples: *Corynebacterium* (26%), *Staphylococcus* (21%), *Moraxella* (15%), *Dolosigranulum* (13%), and *Streptococcus* (5%). Nasopharyngeal microbiome composition varied with age (PERMANOVA, $p < 0.001$, $R^2 = 0.06$); specifically, increasing age was associated with decreases in the CLR-transformed abundances of the bacterial genera *Moraxella* (Spearman's rank correlation; $\rho = -0.40$, $p < 0.0001$) and *Dolosigranulum* ($\rho = -0.32$, $p < 0.0001$) and increases in the CLR-transformed abundances of *Corynebacterium* ($\rho = 0.24$, $p < 0.0001$) and *Staphylococcus* ($\rho = 0.45$, $p < 0.0001$) (**Figure 2**). Two bacterial genera – *Lawsonella* and *Peptoniphilus* – were highly prevalent in participants 12 years of age or older (78% and 69%, respectively), but were identified in only 21% and 19% of children 8 years of age or younger. Nasopharyngeal microbiome composition did not differ significantly in SARS-CoV-2-infected and SARS-CoV-2-uninfected participants (PERMANOVA, $p = 0.10$, $R^2 = 0.004$). However, the composition of the nasopharyngeal microbiome of SARS-CoV-2-infected participants *with* respiratory symptoms differed from that of SARS-CoV-2-infected participants *without* respiratory symptoms (PERMANOVA, $p = 0.002$, $R^2 = 0.009$).

Associations between nasopharyngeal microbiome profile, age, and SARS-CoV-2 status

To further characterize differences in nasopharyngeal microbiome composition by age and SARS-CoV-2 status, we used unsupervised clustering to classify nasopharyngeal samples into seven distinct microbiome profiles (**Figure 3**): *Corynebacterium/Staphylococcus*-dominant (profile 1; n=54, 19%), *Corynebacterium/Dolosigranulum*-dominant (profile 2; n=58, 20%), *Corynebacterium*-dominant (profile 3; n=55, 19%), *Moraxella*-dominant (profile 4; n=44, 16%), *Staphylococcus*-dominant (profile 5; n=44, 15%), *Streptococcus*-dominant (profile 6; n=16, 6%), and *Fusobacterium*-dominant (profile 7; n=10, 4%). Participant age differed markedly by nasopharyngeal microbiome profile (**Table 2**; Kruskal-Wallis test, $p < 0.0001$); however, there were no significant differences in other patient characteristics by microbiome profile. SARS-CoV-2 infection prevalence varied from 69% to 82% by microbiome profile, with the lowest prevalence observed among participants with a *Moraxella*-dominant microbiome profile and the highest prevalence seen among participants with a *Corynebacterium*-dominant microbiome profile. However, in analyses adjusting for age, there were no significant associations between nasopharyngeal microbiome profile and SARS-CoV-2 infection. Among SARS-CoV-2-infected participants, the prevalence of respiratory symptoms varied from 35% to 67% by nasopharyngeal microbiome profile, with the lowest prevalence seen among participants with a *Corynebacterium/Dolosigranulum*-dominant microbiome profile and the highest prevalence observed among participants with *Corynebacterium/Staphylococcus*-dominant or *Corynebacterium*-dominant microbiome profiles. SARS-CoV-2-infected individuals with a *Corynebacterium/Dolosigranulum*-dominant microbiome profile were less likely to have respiratory symptoms than SARS-CoV-2-infected participants with other nasopharyngeal microbiome profiles in age-adjusted analyses (logistic regression; odds ratio: 0.38, 95% confidence interval: 0.18–0.81).

Identification of specific bacterial taxa associated with SARS-CoV-2 infection and COVID-19 symptoms

We next used generalized joint attribute modeling (GJAM) to evaluate associations between specific ASVs and SARS-CoV-2 infection and SARS-CoV-2-associated respiratory symptoms. GJAM allows for the concurrent evaluation of distinct types of data derived from observations of ecological systems, where attributes of the system may be interdependent. Because we observed associations between age and nasopharyngeal microbiome composition, as well as associations between age and SARS-CoV-2 infection and respiratory symptoms, we used GJAM to separately evaluate associations between specific bacterial ASVs and SARS-CoV-2 infection and SARS-CoV-2-associated respiratory symptoms in the context of interactions of these variables with participant age. We identified nine ASVs that were associated with SARS-CoV-2 infection (**Table 3**); for eight of these ASVs, the magnitude of the association varied by participant age. For example, the relative abundance of ASV1163 (*Corynebacterium propinquum*) decreased with increasing participant age and was also higher among SARS-CoV-2-infected participants than uninfected participants independent of age. However, the difference in the relative abundance of ASV1163 between SARS-CoV-2-infected and uninfected participants decreased with age, such that the negative association between the relative abundance of ASV1163 and SARS-CoV-2 infection was primarily observed among young children (**Figure 4A**). We next used GJAM to identify ASVs associated with respiratory symptoms among participants with confirmed SARS-CoV-2 infection. We identified six ASVs that were differentially abundant among SARS-CoV-2-infected participants with respiratory symptoms (**Table 4**); for eight of these ASVs, the magnitude of the association varied by participant age. For example, the relative abundance of ASV336 (*Moraxella lincolnii*) decreased with increasing age and was independently lower among SARS-CoV-2-infected participants *with* respiratory symptoms. However, the difference in the mean relative abundance of ASV336 between SARS-CoV-2-infected participants *with* and *without* respiratory symptoms increased with increasing age, indicating that the negative association

between the relative abundance of this ASV and the presence of SARS-CoV-2-associated respiratory symptoms was observed only in younger age groups (**Figure 4B**).

DISCUSSION

We identified nasopharyngeal microbiome profiles and specific bacterial taxa associated with SARS-CoV-2 infection and the presence of respiratory symptoms in SARS-CoV-2 infected children, adolescents, and young adults. We demonstrated that nasopharyngeal microbiome diversity and composition are strongly associated with age, and that age modifies the associations between specific bacterial taxa and SARS-CoV-2 infection status and the presence of SARS-CoV-2-associated respiratory symptoms. Our findings suggest that the upper respiratory microbiome may be a previously unrecognized and potentially modifiable mechanism by which age influences SARS-CoV-2 susceptibility and respiratory symptoms.

There are accumulating data supporting a key role for the upper respiratory microbiome in the pathogenesis of respiratory virus infections. Prior studies indicate that the upper respiratory microbiome modifies susceptibility to respiratory virus infections, as has been demonstrated in studies of household influenza virus transmission [9]. The upper respiratory microbiome may also influence the symptoms of respiratory virus-infected individuals [21]. For example, among young children with respiratory syncytial virus (RSV), higher abundances of *Haemophilus* species in the nasopharyngeal microbiome are associated with a more exuberant host immune response [10, 22]. Nasopharyngeal microbiome profiles are also associated with inflammatory cytokine levels in nasal wash samples and the development of symptomatic infection in adults following experimental rhinovirus challenge [23]. Further, in animal models, intranasal administration of live bacterial

strains directly modulates immune responses to viral infections [24, 25]. While data from clinical studies are currently lacking, this work suggests that targeted manipulation of the upper respiratory microbiome could be a promising approach to prevent or treat respiratory virus infections.

To date, studies of the upper respiratory microbiome and SARS-CoV-2 infection have primarily been conducted among cohorts of adults presenting with clinical suspicion of COVID-19. De Castilhos and colleagues evaluated the oropharyngeal microbiome in 148 SARS-CoV-2-infected outpatients, 124 hospitalized patients with COVID-19, and 74 healthy adults. They did not identify any significant alterations in microbiome composition between SARS-CoV-2-infected outpatients and healthy controls but found marked dysbiosis among patients hospitalized with severe COVID-19 [26]. Mostafa and colleagues reported lower nasopharyngeal microbial diversity, a lower abundance of the bacterial family Propionibacteriaceae, and a higher abundance of *Corynebacterium accolens* in 40 SARS-CoV-2-infected adults compared to 10 SARS-CoV-2-uninfected adults [27]. In this study, we identified distinct bacterial ASVs associated with SARS-CoV-2 infection and SARS-CoV-2-associated respiratory symptoms. Additionally, we observed a lower prevalence of respiratory symptoms among SARS-CoV-2-infected subjects with *Corynebacterium/Dolosigranulum*-dominated microbiome profiles compared to infected subjects with other microbiome profiles.

Non-*diphtheriae* *Corynebacterium* species and *Dolosigranulum pigrum* were previously shown to have important microbial interactions within the human nasopharynx. *Corynebacterium* abundance within the nasopharyngeal microbiome has been negatively associated with *Streptococcus pneumoniae* colonization among infants and children [28, 29]. Moreover, *Corynebacterium* spp. influence innate immune responses to viral infection in murine models [30, 31]. *D. pigrum* is also generally considered to play a protective role against viral and bacterial infections [32]. Islam and

colleagues recently demonstrated that administration of specific *D. pigrum* strains enhanced resistance to SARS-CoV-2 infection of cultured human respiratory epithelial cells [33]. Further, Smith and colleagues observed decreased abundance of both *Corynebacterium* and *Dolosigranulum* spp. in patients with severe COVID-19 symptoms [34].

Several prior studies evaluated associations between respiratory health and upper respiratory microbiome profiles. Toivonen and colleagues used similar unsupervised clustering methods to create longitudinal nasopharyngeal microbiome profiles for 697 Finnish children during the first two years of life to evaluate associations between microbiome composition, antibiotic exposures, and later asthma development [35]. Despite marked differences in the patient populations and sampling protocols, they identified microbiome profiles dominated by bacterial genera that were also prevalent in our cohort. Teo and colleagues identified similar nasopharyngeal microbiome profiles among Australian infants during the first year of life and at five years of age [36]. Notably, they found that microbiome profiles dominated by *Moraxella*, *Streptococcus*, or *Haemophilus* were more prevalent in samples collected during acute respiratory infections, while profiles dominated by *Staphylococcus*, *Dolosigranulum*, or *Corynebacterium* were more prevalent during periods of health [36]. Finally, Kelly et al. identified five nasopharyngeal microbiome profiles among 319 children less than two years of age in Botswana [29]. In this study, *Streptococcus*-dominant or *Moraxella*-dominant profiles were more common among children with pneumonia or upper respiratory infection symptoms, while the majority of children without respiratory symptoms had microbiome profiles co-dominated by *Corynebacterium* and *Dolosigranulum* [29]. Taken together, these studies demonstrate that the human upper respiratory microbiome is composed largely of species from relatively few bacterial genera, and that microbiome profiles dominated by these species are observed in varied patient populations.

We observed that nasopharyngeal microbiome composition undergoes significant, age-associated shifts during infancy, childhood, and adolescence. Moreover, nearly all the associations between nasopharyngeal microbiome features and SARS-CoV-2 status that we identified in this study were modified by age, indicating that age-associated changes in the nasopharyngeal microbiome likely contribute to the susceptibility to and severity of SARS-CoV-2 infections. Much of our knowledge regarding nasopharyngeal microbiome composition is derived from studies of infants or older adults, and surprisingly little is known about how the upper respiratory microbiome changes during childhood and adolescence [7, 8, 37-39]. Our findings demonstrate that development of the nasopharyngeal microbiome continues throughout childhood and adolescence and highlight the need for future studies to identify the biological or environmental factors that contribute to the shifts in microbiome composition that occur after early childhood.

Our study had several limitations. First, nasopharyngeal samples were collected at a single time point after SARS-CoV-2 exposure; therefore, we were unable to determine if the differences in nasopharyngeal microbiome compositions observed by SARS-CoV-2 infection status preceded, or were the consequence of, SARS-CoV-2 infection. Future studies will need to use longitudinal sampling to evaluate causal relationships between upper respiratory microbiome composition and SARS-CoV-2 infection susceptibility and severity. Secondly, all SARS-CoV-2-infected study participants had relatively mild symptoms; thus, we were unable to identify microbiome features associated with severe COVID-19. Our use of 16S rRNA gene amplicon sequencing prevented us from evaluating other components of the upper respiratory microbiome, including viruses and fungi. Additionally, 16S rRNA gene amplicon experiments have several well-documented biases [40], although we sought to minimize these biases in our study through inclusion of all samples in a single processing run and use of appropriate negative controls. Use of metagenomic sequencing in future studies would enable improved discrimination of bacterial species and analyses of other microbial

kingdoms that are components of the nasopharyngeal microbiome. Finally, residual confounding by unmeasured factors remains possible.

In conclusion, we found that age modifies the associations between specific bacterial taxa and both SARS-CoV-2 infection status and the presence of respiratory symptoms. These findings suggest that development of the nasopharyngeal microbiome during childhood and adolescence may contribute to the differences in SARS-CoV-2 susceptibility and severity observed by age. Future studies should evaluate the potential of the upper respiratory microbiome to serve as a therapeutic target for the prevention and treatment of infections caused by SARS-CoV-2 and other respiratory viruses.

Accepted Manuscript

NOTES

Acknowledgments

We would like to thank the Duke University School of Medicine for use of the Microbiome Core Facility, which performed the DNA extractions and library preparations for this research, and the Duke Sequencing and Genomic Technologies Core Facility, which sequenced these libraries. We offer our sincere gratitude to the children and families who participated in this research.

Availability of data and materials

The sequencing dataset supporting the conclusions of this study is available in the Sequence Read Archive (PRJNA703574). The statistical files and script used for data analyses are also publicly available (https://github.com/alexmcumber/BRAVE_Kids).

Funding

This work was funded by the Duke University School of Medicine and through grants from the Duke Microbiome Center, Children's Miracle Network Hospitals, and the Translating Duke Health Children's Health and Discovery Initiative. AWM was supported by a grant from the National Science Foundation (DGE-1545220) and a grant from the National Institute of Health (T32GM008555). SMH was supported by National Institutes of Health (NIH) training grant (T32-HD094671). MSK was supported by a NIH Career Development Award (K23-AI135090). CWW also reports support for this work from the NIH.

Declarations

MSK reports advisory board fees and consulting fees as an investigator from Adagio Therapeutics, Inc. CWW reports advisory board fees from Roche Molecular Sciences, non-financial support from bioMérieux and Becton Dickinson, a research collaboration with Biofire, and is co-founder of Predigen. AWM reports payments for lectures from Pepperdine University. ATR reports royalties for editorial work from Elsevier, honoraria for participating in a Scientific Advisory Board for Breas US and honoraria for lecturing and development of educational materials for Vapotherm, Inc., and is the unpaid Chair for the Pediatric Section of the Society of Critical Care Medicine (SCCM). NAT reports grants or contracts unrelated to this work from the Centers for Disease Control (CDC Epicenters) and the Rockefeller Foundation, and payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from McGraw-Hill. All other authors have no competing interests to declare.

Accepted Manuscript

REFERENCES

1. Lambert, L. and F.J. Culley, *Innate Immunity to Respiratory Infection in Early Life*. *Front Immunol*, 2017. **8**: p. 1570.
2. Viner, R.M., et al., *Susceptibility to SARS-CoV-2 Infection Among Children and Adolescents Compared With Adults: A Systematic Review and Meta-analysis*. *JAMA Pediatrics*, 2021. **175**(2): p. 143-156.
3. *COVID-19 Stats: COVID-19 Incidence, by Age Group — United States, March 1–November 14, 2020*. *MMWR Morb Mortal Wkly Rep* 2021;69:1664. doi: <http://dx.doi.org/10.15585/mmwr.mm695152a8>.
4. Hurst, J.H., et al., *SARS-CoV-2 Infections Among Children in the Biospecimens from Respiratory Virus-Exposed Kids (BRAVE Kids) Study*. *Clin Infect Dis*, 2020.
5. Castagnoli, R., et al., *Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection in Children and Adolescents: A Systematic Review*. *JAMA Pediatrics*, 2020. **174**(9): p. 882-889.
6. Kim, L., et al., *Hospitalization Rates and Characteristics of Children Aged <18 Years Hospitalized with Laboratory-Confirmed COVID-19 - COVID-NET, 14 States, March 1-July 25, 2020*. *MMWR Morb Mortal Wkly Rep*, 2020. **69**(32): p. 1081-1088.
7. Mika, M., et al., *Dynamics of the nasal microbiota in infancy: A prospective cohort study*. *Journal of Allergy and Clinical Immunology*, 2015. **135**(4): p. 905-912. e11.
8. Biesbroek, G., et al., *Early respiratory microbiota composition determines bacterial succession patterns and respiratory health in children*. *American journal of respiratory and critical care medicine*, 2014. **190**(11): p. 1283-1292.

9. Tsang, T.K., et al., *Association between the respiratory microbiome and susceptibility to influenza virus infection*. *Clinical Infectious Diseases*, 2020. **71**(5): p. 1195-1203.
10. de Steenhuijsen Piters, W.A., et al., *Nasopharyngeal Microbiota, Host Transcriptome and Disease Severity in Children with Respiratory Syncytial Virus Infection*. *American journal of respiratory and critical care medicine*, 2016(ja).
11. Gilbert, J.A., et al., *Meeting report: the terabase metagenomics workshop and the vision of an Earth microbiome project*. *Stand Genomic Sci*, 2010. **3**(3): p. 243-8.
12. Callahan, B.J., et al., *DADA2: high-resolution sample inference from Illumina amplicon data*. *Nature methods*, 2016. **13**(7): p. 581.
13. Escapa, I., et al., *Construction of habitat-specific training sets to achieve species-level assignment in 16S rRNA gene datasets*. *Microbiome*, 2020. **8**: p. 1-16.
14. Davis, N.M., et al., *Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data*. *Microbiome*, 2018. **6**(1): p. 226.
15. McMurdie, P.J. and S. Holmes, *phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data*. *PloS one*, 2013. **8**(4): p. e61217.
16. Lahti, L. and S. Shetty, *microbiome R package*. 2017.
17. Gloor, G.B., et al., *It's all relative: analyzing microbiome data as compositions*. *Annals of epidemiology*, 2016. **26**(5): p. 322-329.
18. Oksanen, J., et al., *The vegan package*. *Community ecology package*, 2007. **10**(631-637): p. 719.
19. Clark, J.S., et al., *Generalized joint attribute modeling for biodiversity analysis: Median-zero, multivariate, multifarious data*. *Ecological Monographs*, 2017. **87**(1): p. 34-56.

20. R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
21. Pichon, M., B. Lina, and L. Josset, *Impact of the respiratory microbiome on host responses to respiratory viral infection*. *Vaccines*, 2017. **5**(4): p. 40.
22. Ederveen, T.H., et al., *Haemophilus is overrepresented in the nasopharynx of infants hospitalized with RSV infection and associated with increased viral load and enhanced mucosal CXCL8 responses*. *Microbiome*, 2018. **6**(1): p. 1-13.
23. Allen, E.K., et al., *Characterization of the nasopharyngeal microbiota in health and during rhinovirus challenge*. *Microbiome*, 2014. **2**(1): p. 22.
24. Tomosada, Y., et al., *Nasally administered Lactobacillus rhamnosus strains differentially modulate respiratory antiviral immune responses and induce protection against respiratory syncytial virus infection*. *BMC immunology*, 2013. **14**(1): p. 40.
25. Zelaya, H., et al., *Nasal priming with immunobiotic Lactobacillus rhamnosus modulates inflammation–coagulation interactions and reduces influenza virus-associated pulmonary damage*. *Inflammation Research*, 2015. **64**(8): p. 589-602.
26. de Castilhos, J., et al., *Severe dysbiosis and specific Haemophilus and Neisseria signatures as hallmarks of the oropharyngeal microbiome in critically ill COVID-19 patients*. *Clin Infect Dis*, 2021.
27. Mostafa, H.H., et al., *Metagenomic next-generation sequencing of nasopharyngeal specimens collected from confirmed and suspect covid-19 patients*. *Mbio*, 2020. **11**(6).
28. Khamash, D.F., et al., *The Association Between the Developing Nasal Microbiota of Hospitalized Neonates and Staphylococcus aureus Colonization*. *Open Forum Infectious Diseases*, 2019. **6**(4).

29. Kelly, M.S., et al., *The nasopharyngeal microbiota of children with respiratory infections in Botswana*. The Pediatric infectious disease journal, 2017. **36**(9): p. e211.
30. Kanmani, P., et al., *Respiratory commensal bacteria Corynebacterium pseudodiphtheriticum improves resistance of infant mice to respiratory syncytial virus and Streptococcus pneumoniae superinfection*. Frontiers in microbiology, 2017. **8**: p. 1613.
31. Mak, N., E. Schiltknecht, and G. Ada, *Protection of mice against influenza virus infection: enhancement of nonspecific cellular responses by Corynebacterium parvum*. Cellular immunology, 1983. **78**(2): p. 314-325.
32. Man, W.H., W.A. de Steenhuijsen Piters, and D. Bogaert, *The microbiota of the respiratory tract: gatekeeper to respiratory health*. Nat Rev Microbiol, 2017. **15**(5): p. 259-270.
33. Islam, M.A., et al., *Modulates Immunity against SARS-CoV-2 in Respiratory Epithelial Cells*. Pathogens, 2021. **10**(6).
34. Smith, N., et al., *Distinct systemic and mucosal immune responses during acute SARS-CoV-2 infection*. Nat Immunol, 2021. **22**(11): p. 1428-1439.
35. Toivonen, L., et al., *Antibiotic Treatments During Infancy, Changes in Nasal Microbiota, and Asthma Development: Population-based Cohort Study*. Clinical Infectious Diseases, 2020. **72**(9): p. 1546-1554.
36. Teo, S.M., et al., *The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development*. Cell host & microbe, 2015. **17**(5): p. 704-715.
37. Bogaert, D., et al., *Variability and diversity of nasopharyngeal microbiota in children: a metagenomic analysis*. Plos one, 2011. **6**(2): p. e17035.

38. Hahn, A., et al., *Microbial diversity within the airway microbiome in chronic pediatric lung diseases*. *Infection, Genetics and Evolution*, 2018. **63**: p. 316-325.
39. Pérez-Losada, M., et al., *Nasopharyngeal Microbiome Diversity Changes over Time in Children with Asthma*. *PLoS One*, 2017. **12**(1): p. e0170543.
40. Brooks, J.P., et al., *The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies*. *BMC Microbiology*, 2015. **15**(1): p. 66.

Accepted Manuscript

Figure Legends

Figure 1. Nasopharyngeal microbiome alpha diversity by age. Shannon diversity (A) and the number of unique amplicon sequence variants (B) are shown by participant age. Each point represents an individual sample and lines correspond to the fit of the linear model between age and each alpha diversity measure. ASV, amplicon sequence variant

Figure 2. Relative abundances of highly abundant bacterial genera by age. Each bar depicts the mean relative abundances of highly abundant genera in nasopharyngeal samples from participants in a specific age category. Only the nine most highly abundant genera within nasopharyngeal samples from the entire study population are shown. Age is shown as a categorical variable only for graphical representation; all statistical analyses included age as a continuous variable.

Figure 3. Nasopharyngeal microbiome profiles identified by unsupervised clustering. A) Principal coordinate plot of Euclidean distances demonstrating clustering of nasopharyngeal samples by microbiome profile. Each dot corresponds to a single nasopharyngeal sample. Centroids are shown as the confluence of the lines arising from individual points from each microbiome profile. Ellipses define the regions containing 95% of all samples that can be drawn from the underlying multivariate t distribution. B) Each bar depicts the mean relative abundances of highly abundant genera in nasopharyngeal samples assigned to specific microbiome profiles. Only the nine most highly abundant genera within nasopharyngeal samples from the entire study population are shown.

Figure 4. Interactive relationships between participant age, the relative abundances of specific bacterial ASVs in the nasopharyngeal microbiome, and SARS-CoV-2 status. A) Bar chart depicting

differences in the mean relative abundance of ASV1163 (*Corynebacterium propinquum*) among SARS-CoV-2-infected participants relative to uninfected participants in different age categories. The gold line was constructed using the GJAM estimates for the association of SARS-CoV-2 infection with the relative abundance of ASV1163 (intercept) and the association of the interaction term between SARS-CoV-2 infection and age with the relative abundance of ASV1163 (slope). Higher relative abundances of ASV1163 were observed in SARS-CoV-2-infected compared to uninfected participants across all ages but these differences were more pronounced in young children. B) Differences in mean relative abundance of ASV336 (*Moraxella lincolnii*) between SARS-CoV-2-infected participants *with* respiratory symptoms and SARS-CoV-2-infected participants *without* respiratory symptoms are depicted by age category. Dark (light) grey bars represent age categories in which ASV336 was more (less) abundant among SARS-CoV-2-infected participants *with* respiratory symptoms compared to SARS-CoV-2-infected participants without respiratory symptoms. The purple line was constructed using the GJAM estimates for the association of SARS-CoV-2-associated respiratory symptoms with the relative abundance of ASV336 (intercept) and the association of the interaction term between respiratory symptoms and age with the relative abundance of ASV336 (slope). The difference in the mean relative abundance of ASV336 between SARS-CoV-2-infected participants *with* and *without* respiratory symptoms differed by age, such that this ASV was less abundant in the context of SARS-CoV-2-associated respiratory symptoms among young children and more abundant in the context of SARS-CoV-2-associated respiratory symptoms in older age groups. Lines were fit using the regression coefficients generated using GJAM. Age is shown as a categorical variable only for graphical representation; all statistical analyses included age as a continuous variable.

Table 1. Characteristics of the study population

	SARS-CoV-2-Exposed, Uninfected (n=74)		SARS-CoV-2-Infected <i>Without</i> Respiratory Symptoms (n=98)		SARS-CoV-2-Infected <i>With</i> Respiratory Symptoms (n=113)		<i>p</i> ^a
	N (or median)	% (or IQR)	N (or median)	% (or IQR)	N (or median)	% (or IQR)	
Age, years	9.5	(5.1–15.8)	9.3	(4.8–13.2)	14.1	(6.3–17.5)	0.01
Female sex	37	50%	53	54%	59	52%	0.87
Race							<0.0001
Black or African-American	2	3%	6	6%	8	7%	
Latino or Hispanic-American	42	57%	86	88%	100	88%	
Non-Hispanic white	30	41%	6	6%	5	4%	
Comorbidities ^b							
Asthma	8	11%	6	6%	9	8%	0.53
Obesity (BMI ≥95 th percentile for age)	20	27%	24	24%	40	35%	0.19
Environmental tobacco smoke in home	11	15%	9	9%	14	12%	0.51
Receipt of antibiotic in prior 30 days	1	1%	2	2%	3	3%	>0.99
Receipt of probiotic in prior 30 days	3	4%	0	0%	1	1%	0.09

IQR, interquartile range; BMI, body mass index

^b*p* values were estimated using Chi square or Fisher's exact tests for categorical variables and Kruskal-Wallis tests for continuous variables

^bOther comorbidities included hypertension (n=5), congenital heart disease (n=3), chronic neurological disorder (n=3), chronic kidney disease (n=2), and malignancy (n=1)

Table 2. Characteristics of study participants and nasopharyngeal microbial communities by microbiome profile

	Nasopharyngeal Microbiome Profile ^a							<i>p</i> ^b
	1 (n=54)	2 (n=58)	3 (n=55)	4 (n=48)	5 (n=44)	6 (n=16)	7 (n=10)	
Age (years), median (IQR)	16.5 (12.9–)	7.6 (3.9–12.1)	15.9 (10.6–)	3.8 (2.0–7.2)	9.8 (8.3–13.2)	7.7 (1.0–11.5)	7.0 (6.0–13.4)	<0.0001
Female sex	29 (53%)	29 (50%)	33 (60%)	31 (65%)	17 (39%)	7 (44%)	3 (30%)	0.12
Race/ethnicity								
Black or African-American	5 (9%)	4 (7.0%)	4 (7%)	1 (2%)	1 (2%)	1 (6%)	0 (0%)	0.23
Latino or Hispanic-American	37 (69%)	46 (79%)	46 (84%)	38 (79%)	40 (91%)	11 (69%)	10 (100%)	
Non-Hispanic white	12 (22%)	8 (14%)	5 (9%)	9 (19%)	3 (7%)	4 (25%)	0 (0%)	
Comorbidities								
Asthma	5 (9%)	2 (3%)	5 (9%)	3 (6.3%)	8 (18%)	0 (0%)	0 (0%)	0.19
Obesity (BMI ≥95 th percentile for age)	20 (37%)	15 (26%)	18 (33%)	13 (27%)	11 (25%)	3 (19%)	4 (40%)	0.67
Environmental tobacco smoke in	6 (11%)	9 (16%)	9 (16%)	4 (8.3%)	3 (7%)	3 (19%)	0 (0%)	0.52
Receipt of antibiotic in prior 30	2 (4%)	1 (2%)	1 (2%)	1 (2%)	1 (2%)	0 (0%)	0 (0%)	0.97
Receipt of probiotic in prior 30	3 (6%)	0 (0%)	0 (0%)	0 (0%)	1 (2%)	0 (0%)	0 (0%)	0.18
SARS-CoV-2 infection	39 (72%)	43 (74%)	45 (82%)	33 (69%)	32 (72%)	11 (69%)	8 (80%)	0.80
with respiratory symptoms	26 (67%)	15 (35%)	30 (67%)	18 (55%)	13 (41%)	6 (55%)	5 (63%)	0.03
without respiratory symptoms	13 (33%)	28 (65%)	15 (33%)	15 (45%)	19 (59%)	5 (45%)	3 (38%)	
Shannon diversity index, median (IQR)	1.64 (1.37–)	1.15 (1.05–)	1.04 (0.91–)	0.91 (0.81–)	0.96 (0.45–)	1.5 (1.22–)	1.72 (1.37–)	<0.0001
Number of unique ASVs, median (IQR)	72 (45–92)	66 (45–83)	65 (49–82)	58 (42–78)	71 (53–101)	76 (59–91)	92 (57–106)	0.16

^aCorresponding microbiome profiles are: 1=*Corynebacterium/Staphylococcus*-dominant, 2=*Corynebacterium/Dolosigranulum*-dominant, 3=*Corynebacterium*-dominant, 4=*Moraxella*-dominant, 5=*Staphylococcus*-dominant, 6=*Streptococcus*-dominant, 7=*Fusobacterium*-dominant

^b*p* values were estimated using Fisher's exact tests for categorical variables and Kruskal-Wallis tests for continuous variables

IQR, interquartile range; BMI, body mass index; ASV, amplicon sequence variant

Table 3. Differentially abundant bacterial ASVs in SARS-CoV-2-infected and uninfected participants in GJAM analyses

ASV	Bacterial species	Age		SARS-CoV-2 Infection		Age * SARS-CoV-2 Infection	
		Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
306	<i>Moraxella nonliquefaciens</i>	-0.020	(-0.023, -0.016)	-0.108	(-0.158, -0.055)	0.005	(0.001, 0.009)
629	<i>Prevotella nanceiensis</i>	-0.002	(-0.003, -0.001)	-0.022	(-0.037, -0.006)	0.002	(0.0003, 0.003)
692	<i>Prevotella intermedia</i>	-0.002	(-0.003, -0.0004)	-0.020	(-0.038, -0.001)	NS	NS
712	<i>Prevotella melaninogenica</i>	-0.002	(-0.003, -0.0009)	-0.016	(-0.032, -0.0006)	0.001	(0.0007, 0.0002)
1095	<i>Corynebacterium tuberculostearicum</i>	0.006	(0.004, 0.007)	0.022	(0.002, 0.041)	-0.003	(-0.004, -0.001)
1163	<i>Corynebacterium propinquum</i>	-0.004	(-0.007, -0.0002)	0.090	(0.040, 0.142)	-0.004	(-0.009, -0.0004)
1165	<i>Corynebacterium propinquum</i>	NS	NS	-0.033	(-0.059, -0.007)	0.003	(0.001, 0.006)
1488	<i>Streptococcus mitis</i>	-0.002	(-0.004, -0.001)	-0.202	(-0.038, -0.002)	NS	NS
1581	<i>Gemella morbillorum</i>	-0.003	(-0.004, -0.001)	-0.022	(-0.040, -0.003)	0.002	(0.0003, 0.003)

ASV, amplicon sequence variant; CI, confidence interval; NS, not significant

Table 4. Differentially abundant bacterial ASVs in SARS-CoV-2-infected participants *with* and *without* respiratory symptoms in GJAM analyses

ASV	Bacterial Species	Age		SARS-CoV-2 Respiratory Symptoms		Age * SARS-CoV-2 Respiratory Symptoms	
		Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
336	<i>Moraxella lincolnii</i>	-0.006	(-0.008, -0.004)	-0.081	(-0.111, -0.051)	0.007	(0.004, 0.009)
339	<i>Moraxella lincolnii</i>	-0.002	(-0.003, -0.0008)	-0.018	(-0.032, -0.004)	0.002	(0.0003, 0.003)
692	<i>Prevotella intermedia</i>	NS	NS	0.024	(0.008, 0.039)	-0.002	(-0.003, -0.0002)
1283	<i>Fusobacterium nucleatum</i>	0.003	(0.0009, 0.005)	0.038	(0.008, 0.067)	-0.004	(-0.007, -0.002)
1519	<i>Streptococcus pyogenes</i>	NS	NS	0.043	(0.002, 0.071)	NS	NS
2155	<i>Mycoplasma lipophilum</i>	-0.001	(-0.002, -0.0003)	-0.014	(-0.027, -0.001)	NS	NS

ASV, amplicon sequence variant; CI, confidence interval; NS, not significant

Figure 1

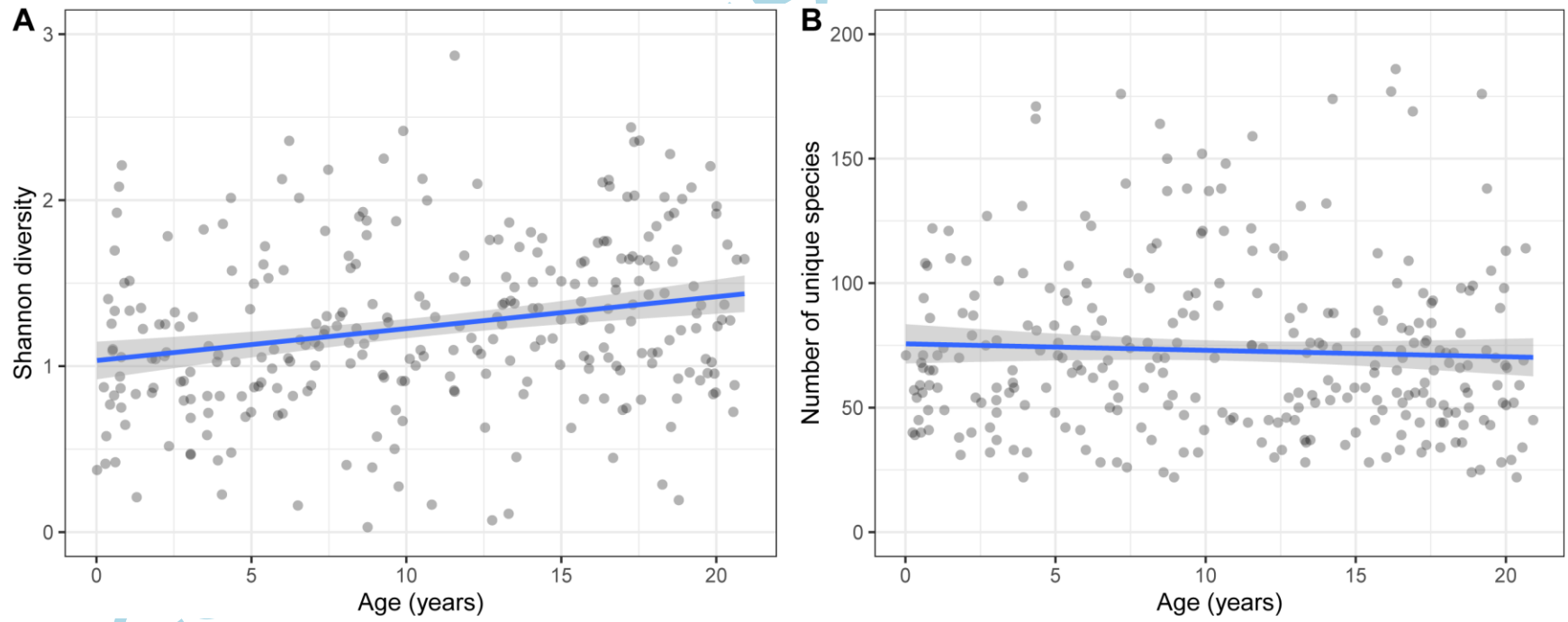


Figure 2

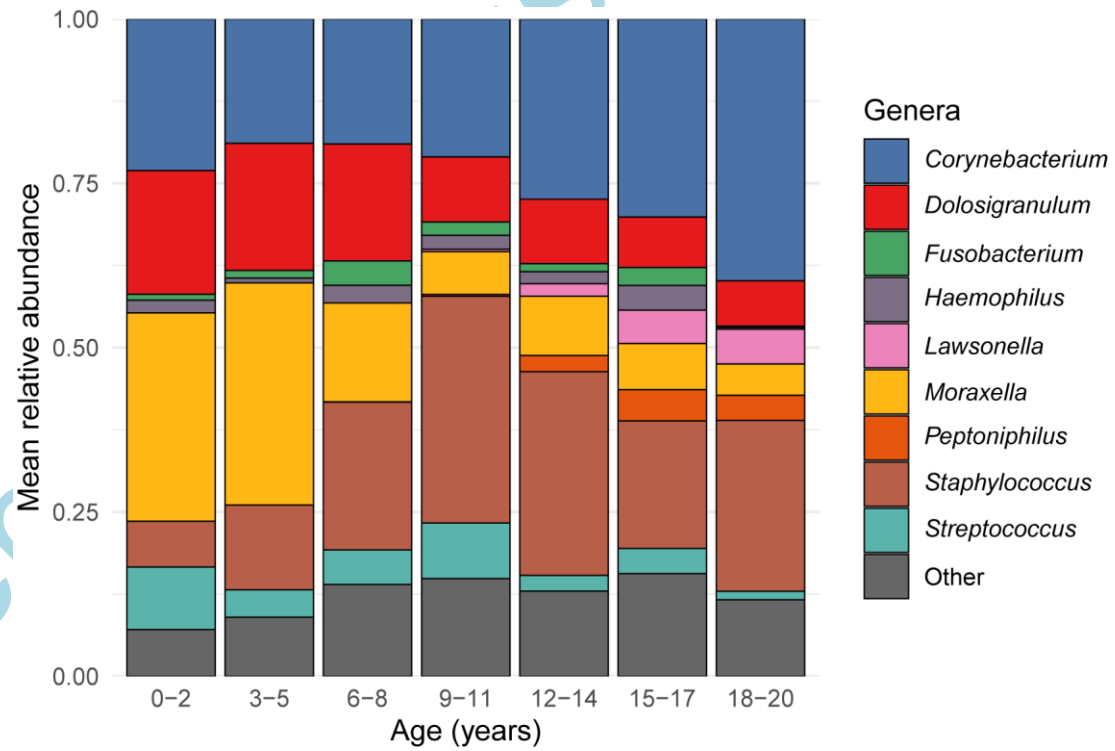


Figure 3

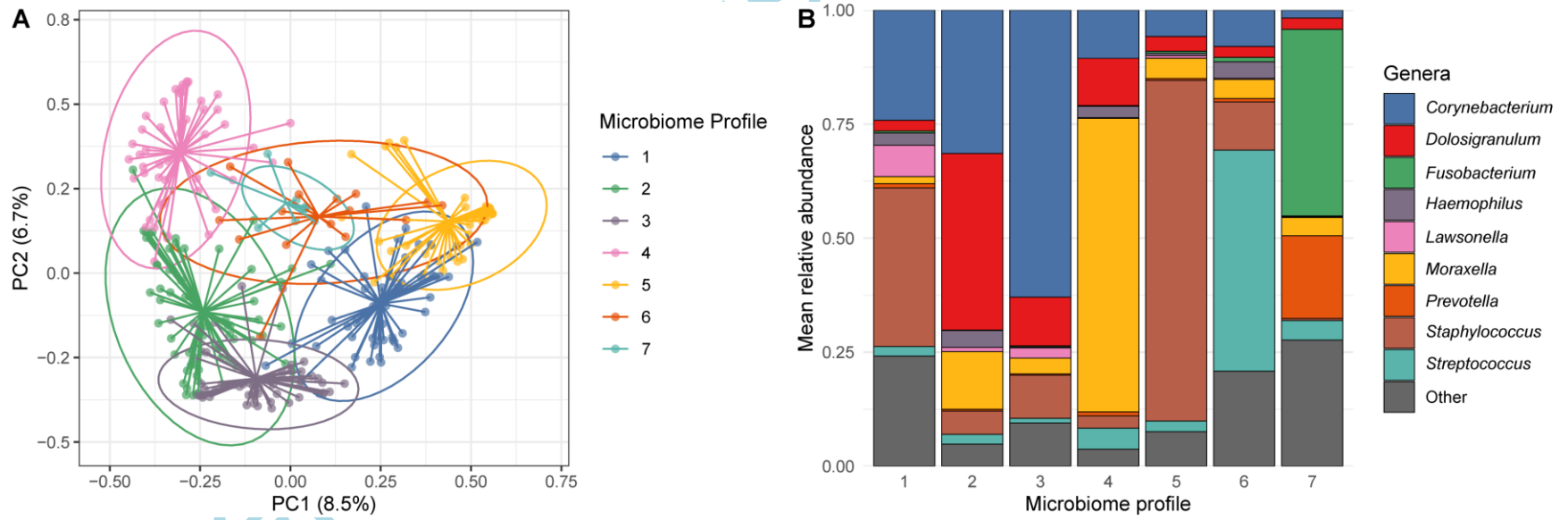
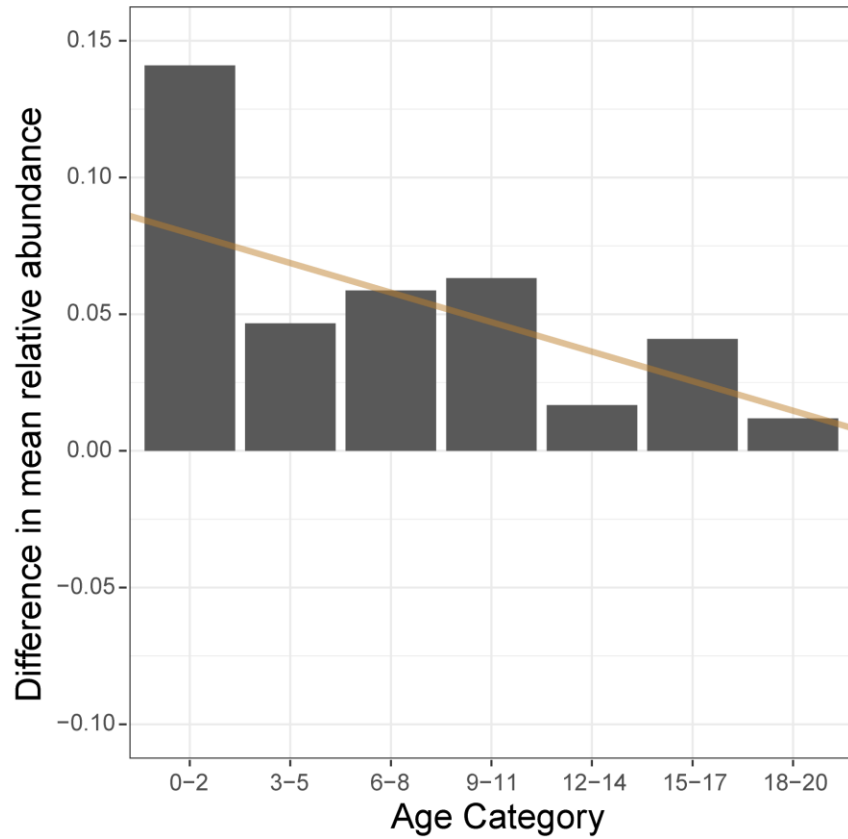


Figure 4

A ASV1163 (*Corynebacterium propinquum*)



B ASV336 (*Moraxella lincolnii*)

