JCI insight

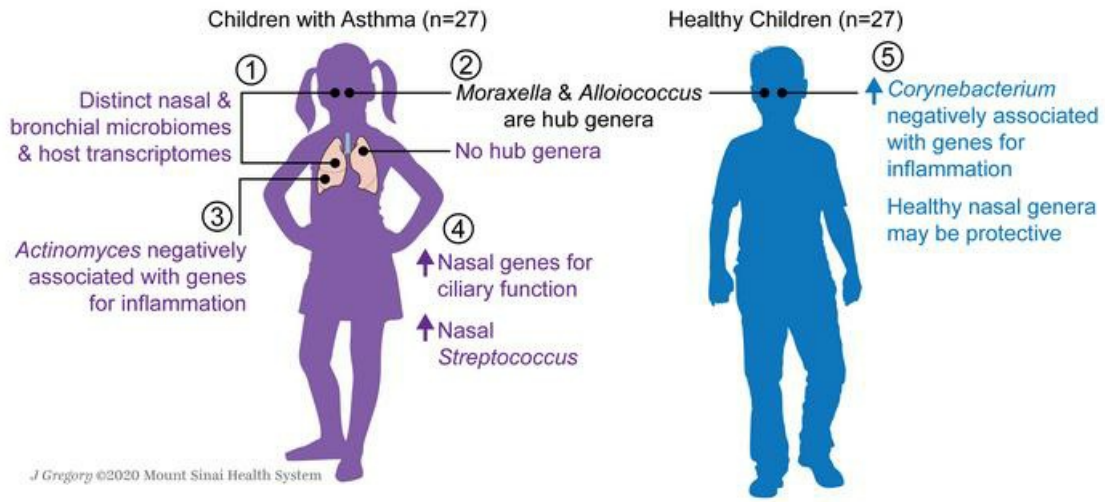
Integrative study of the upper and lower airway microbiome and transcriptome in asthma

Yoojin Chun, … , Eric Schadt, Supinda Bunyavanich

JCI Insight. 2020;5(5):e133707. https://doi.org/10.1172/jci.insight.133707.

Research Article | Microbiology | Pulmonology

Graphical abstract



Find the latest version:

https://jci.me/133707/pdf

JCI insight

RESEARCH ARTICLE

Integrative study of the upper and lower airway microbiome and transcriptome in asthma

Yoojin Chun,[1] Anh Do,[1] Galina Grishina,[2] Alexander Grishin,[2] Gang Fang,[1] Samantha Rose,[3] Chantal Spencer,[3] Alfin Vicencio,[3] Eric Schadt,[1] and Supinda Bunyavanich[1,2]

[1]Institute for Data Science and Genomic Technology, Department of Genetics and Genomic Sciences, [2]Division of Allergy and Immunology, Department of Pediatrics, and [3]Division of Pulmonary Medicine, Department of Pediatrics, Icahn School of Medicine at Mount Sinai, New York, New York, USA.

Relatively little is known about interactions between the airway microbiome and airway host transcriptome in asthma. Since asthma affects and is affected by the entire airway, studying the upper (e.g., nasal) and lower (e.g., bronchial) airways together represents a powerful approach to understanding asthma. Here, we performed a systematic, integrative study of the nasal and bronchial microbiomes and nasal and bronchial host transcriptomes of children with severe persistent asthma and healthy controls. We found that (a) the microbiomes and host transcriptomes of asthmatic children are each distinct by site (nasal versus bronchial); (b) among asthmatic children, *Moraxella* and *Alloiococcus* are hub genera in the nasal microbiome, while there are no hubs among bronchial genera; (c) bronchial *Actinomyces* is negatively associated with bronchial genes for inflammation, suggesting *Actinomyces* may be protective; (d) compared with healthy children, asthmatic children express more nasal genes for ciliary function and harbor more nasal *Streptococcus*; and (e) nasal genera such as *Corynebacterium* are negatively associated with significantly more nasal genes for inflammation in healthy versus asthmatic children, suggesting a potentially stronger protective role for such nasal genera in healthy versus asthmatic children. Our systematic, integrative study provides a window into host-microbiome associations in asthma.

Introduction

Asthma is a prevalent disease affecting millions of people worldwide of all ages and ethnic backgrounds and is the most common chronic respiratory disease in children (1). Among individuals with asthma, those with severe persistent asthma have the most impaired health and quality of life, requiring chronic high-dose medications (2) and living at risk for life-threatening respiratory failure, recurrent hospitalizations, and high medical expenses (3). Although research on severe persistent asthma has been carried out to better understand this complex disease, it is still suboptimally understood (4). There has been increasing interest in how the airway microbiome may play a role in asthma (5–15), as well as interest in deciphering pathophysiologic processes and endotypes underlying asthma via studies of the airway host transcriptome (5, 16–18). However, airway microbiome and airway transcriptome studies are commonly conducted independently; thus, relatively little is known about interactions between the airway microbiome and airway transcriptome in asthma (5).

Since asthma affects and is affected by the entire airway, studying the upper (e.g., nasal) and lower (e.g., bronchial) airways together represents a powerful approach to understanding asthma (19). However, such studies can be challenging to conduct, particularly in children, due to the limitations of sampling the bronchial airway, which requires invasive bronchoscopy and anesthesia. While there have been a few prior studies of children, which have examined the microbiome of the nasal and bronchial airways together (9, 20), these were not studies of asthma; instead, they were opportunistic studies of children undergoing bronchoscopy for a variety of heterogeneous indications. Therefore, they did not yield insight on the airway dynamics of asthma in particular. To our knowledge, there has not been a study that has investigated the microbiomes and host transcriptomes of the nasal and bronchial airways of children with asthma. In this study, we systematically profiled and analyzed the nasal and bronchial microbiomes and nasal and bronchial transcriptomes of children with severe persistent asthma and healthy controls (Figure 1) to identify nasal

Conflict of interest: The authors have declared that no conflict of interest exists.

Copyright: © 2020, American Society for Clinical Investigation.

Submitted: September 20, 2019
Accepted: February 13, 2020
Published: March 12, 2020.

Reference information: *JCI Insight*. 2020;5(5):e133707. https://doi.org/10.1172/jci.insight.133707.

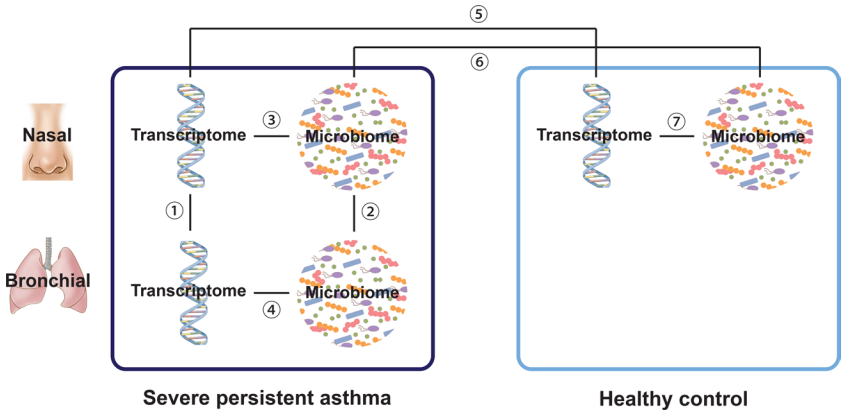


Figure 1. Study design and analytic flow. For each child with severe persistent asthma (n = 27), 4 matched samples were collected and profiled, including nasal transcriptome, nasal microbiome, bronchial transcriptome, and bronchial microbiome. For each healthy control (i.e., child without asthma, n = 27), samples for nasal transcriptome and nasal microbiome were collected and profiled; bronchial samples were not collected because bronchoscopy is not indicated in healthy children. The circled numbers, referred to as steps, represent the analytic steps taken in this systematic study. Among children with severe persistent asthma, we compared the nasal versus bronchial transcriptome (step 1), and the nasal versus bronchial microbiome (step 2). We also assessed positive and negative associations between genera abundances among nasal microbiota, among bronchial microbiota, and between nasal and bronchial microbiota in children with severe persistent asthma (step 2). Next, using the matched samples from these children with severe persistent asthma, we characterized associations between nasal transcriptome and nasal microbiota (step 3), as well as associations between bronchial transcriptome and bronchial microbiota (step 4). We then compared the nasal transcriptome of children with severe persistent asthma with that of healthy controls (step 5), as well as the nasal microbiome of children with severe persistent asthma with that of healthy controls (step 6). Finally, we characterized the associations between nasal transcriptome and nasal microbiota in healthy controls (step 7).

and bronchial microbiota, host nasal and bronchial gene transcripts, and their associations that characterize severe persistent asthma.

Results

Study population, sample collection, and sample profiling. We recruited 27 children with severe persistent asthma (median age = 11 years, IQR = 8 years) and 27 healthy control subjects of similar age (median age = 13 years, IQR = 6 years) (Supplemental Table 1; supplemental material available online with this article; https://doi.org/10.1172/jci.insight.133707DS1). Among the children with severe persistent asthma, 26 (96%) had a history of an emergency room visit ever, including 18 (67%) who had been to the emergency room in the past year. Nineteen (70%) had been previously hospitalized for asthma, including 8 (30%) in the past year. The median Asthma Control Test (ACT) score of children with severe persistent asthma was 13 (IQR = 7.5), while all of the healthy children uniformly scored the highest possible score (ACT = 25) for asthma control.

Matched samples were obtained from all subjects (Figure 1). Children with severe persistent asthma underwent bronchoscopy for their clinical care, during which the following 4 matched samples were collected: nasal and bronchial brushings for transcriptome profiling, and nasal swabs and bronchoalveolar lavage (BAL) fluid for microbiome profiling. Healthy children did not undergo bronchoscopy due to the invasiveness of this procedure for healthy children. Matched nasal brushings and nasal swabs were collected from all healthy children. RNA-sequencing (RNA-seq) was performed on RNA isolated from the nasal and bronchial brushing samples, and 16S rRNA-seq was performed on DNA isolated from nasal swabs and BAL fluid. Following data processing and quality control, there were RNA-seq data (n = 17,337 genes) and 16S sequencing data (1,867,809 reads clustering into 3052 OTUs) from the samples for analysis.

Distinct nasal and bronchial transcriptomes in children with severe persistent asthma. We first compared transcriptome profiles from the nasal and bronchial airways of children with severe persistent asthma (step 1 in Figure 1). Along the first 2 principal components computed on the transcriptome data for each anatomic site, the samples clustered by anatomic site (Figure 2A), with 4522 genes differentially expressed (DE) between the nasal and bronchial airways with threshold of FDR ≤ 0.05 and $|\log_2$ fold change$|$ > 1 (Figure 2B). To characterize the functional differences represented by this DE signature, we searched

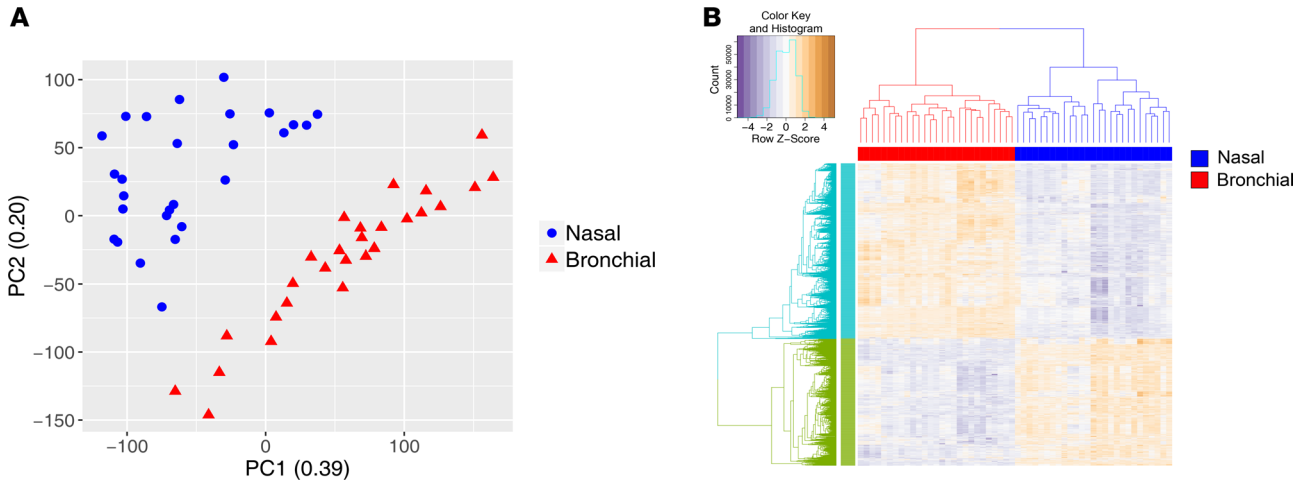


Figure 2. Nasal versus bronchial transcriptome in children with severe persistent asthma (step 1 in Figure 1). Paired nasal and bronchial samples were collected from children with severe persistent asthma (n = 27), and RNA-seq was used for transcriptome profiling. (**A**) Principal Component Analysis (PCA) plot of nasal and bronchial transcriptomes of children with severe persistent asthma shows distinct clustering by anatomic site. Each point represents a sample. (**B**) Heatmap of genes differentially expressed (FDR ≤ 0.05, and $|\log_2$ fold change$|$ > 1) between nasal and bronchial transcriptomes in children with severe asthma shows distinct clusters. Rows show the 2618 genes upregulated in the bronchial versus nasal transcriptome (cyan), and the 1904 genes upregulated in the nasal versus bronchial transcriptome (green). Columns show the 27 bronchial (red) and 27 nasal (blue) samples.

for biological processes significantly enriched (FDR ≤ 0.05) in this signature using Gene Ontology (GO) enrichment analysis (21, 22). The top 3 GO terms for genes upregulated in the bronchial brushings of these children with asthma, ranked by FDR, were cilium assembly (fold enrichment = 5.0, FDR = 1.76 × 10^{-26}), cilium morphogenesis (fold enrichment = 4.7, FDR = 2.16 × 10^{-25}), and cilium movement (fold enrichment = 7.5, FDR = 1.58 × 10^{-11}). In contrast, the top GO terms for genes upregulated in nasal brushings from these asthmatic children were inflammatory response (fold enrichment = 2.2, FDR = 3.31 × 10^{-7}) and epidermis development (fold enrichment = 3.5, FDR = 3.19 × 10^{-5}). We provide the full list of DE genes in Supplemental Table 2 and the full list of GO processes significantly enriched for these DE genes in Supplemental Table 3. Sensitivity analyses showed no DE genes among nasal or among bronchial samples from asthmatics based on ACT score or medication use (Supplemental Table 4).

Distinct nasal and bronchial microbiomes in children with severe persistent asthma. Next, we analyzed the nasal and bronchial microbiomes in children with severe persistent asthma (step 2 in Figure 1). Species diversity within site (α diversity) by Shannon diversity index was higher in the bronchial versus nasal airway (FDR = 3.54 × 10^{-5}) (Figure 3A). Community composition differences between sites (β diversity) was also significantly different between the nasal and bronchial airways based on both weighted UniFrac distance (1-way PERMANOVA pseudo F = 38.8, P = 1.0 × 10^{-3}) (Figure 3B) and unweighted UniFrac distance (1-way PERMANOVA pseudo F = 11.0, P = 1.0 × 10^{-3}) (Supplemental Figure 1). Many genera were differentially abundant between the nasal and bronchial airways in children with severe asthma (Figure 3C). *Corynebacterium* (mean = 21.6%, IQR = 25.2%), *Staphylococcus* (mean = 20.9%, IQR = 31.4%), and *Moraxella* (mean = 20.8%, IQR = 23.2%) were the top prevalent genera in the nasal airway, whereas *Veillonella* (mean = 19.4%, IQR = 12.8%), *Prevotella* (mean = 16.4%, IQR = 13.4%), *Streptococcus* (mean = 11.0%, IQR = 8.5%), and *Neisseria* (mean = 10.7%, IQR = 8.5%) were the most common genera in the bronchial airway (Figure 3C). Sensitivity analyses showed no differences in α diversity among nasal or among bronchial samples from asthmatics based on ACT score or medication use (Supplemental Table 4).

Denser microbial associations in the nasal versus bronchial airway of severe persistent asthmatics. To characterize relationships between microbiota within the nasal airway, within the bronchial airway, and between the nasal and bronchial airways (step 2 in Figure 1), we constructed respective networks based on an ensemble of 4 statistics (23) derived from permutation-renormalization and bootstrapping of 2 correlation measures (Pearson, Spearman) and 2 dissimilarity measures (Bray-Curtis dissimilarity and Kullback-Leibler divergence; refs. 24, 25). We refer to this network approach as the ensemble4 method (see Methods). Summary characteristics of the microbiome networks constructed are shown in Supplemental Table 5.

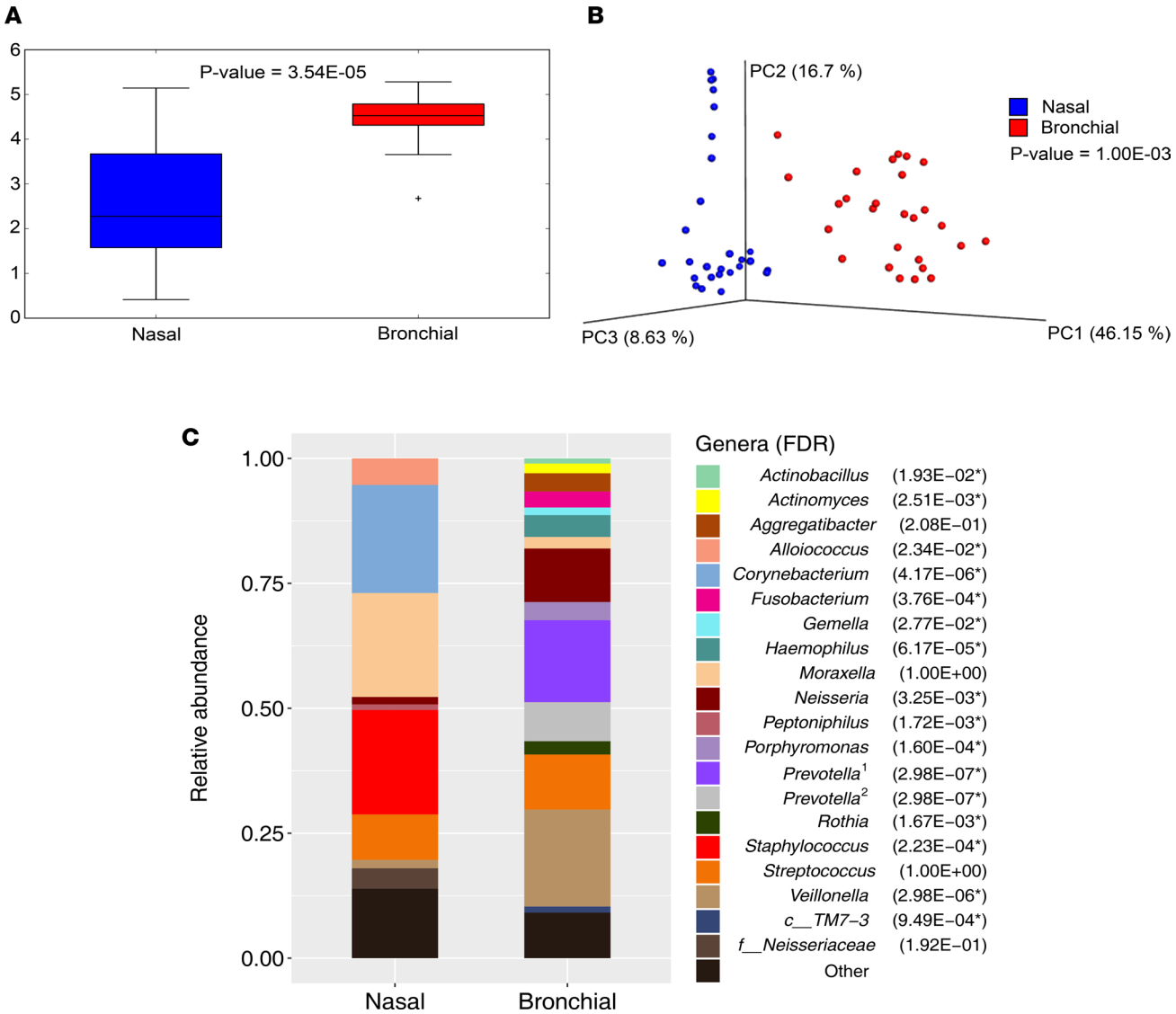


Figure 3. Nasal versus bronchial microbiome in children with severe persistent asthma (step 2 in Figure 1). (**A**) Increased bacterial diversity of bronchial versus nasal microbiome in children with severe persistent asthma (n = 27). The α diversity was estimated by Shannon diversity index by subsampling 10 times for each sample at a rarefaction depth of 1950. P values were calculated using Wilcoxon signed-rank test and FDR correction. (**B**) Principal Coordinates Analysis (PCoA) plot of weighted UniFrac distance shows clear separation of bacterial composition between the nasal and bronchial microbiome of children with severe persistent asthma (n = 27). Each point represents a sample and is colored by airway site. P values were calculated using PERMANOVA with 1000 permutations. (**C**) Distinct mean relative abundances of bacterial genera in the nasal versus bronchial microbiomes of children with severe persistent asthma (n = 27). Genera representing < 1% of total abundance in each site are shown as Other. Sixteen of the 20 genera (80%) were differentially abundant by site (Wilcoxon signed-rank test FDR ≤ 0.05) and are marked with asterisks. For unnamed genera, the closest higher level named taxon is indicated. [1]*Prevotella* genus affiliated with the *Prevotellaceae* family. [2]*Prevotella* genus proposed as affiliated with *Paraprevotellaceae* by the Greengenes database v.13.8.

The network of associations among microbial genera in the nasal airway of children with severe persistent asthma is shown in Figure 4A, and the list of all associations in the network is shown in Supplemental Table 6. The mean number of associations for each genus in the network was 5.9, with *Moraxella* and *Alloiococcus* exhibiting the highest number of associations with other genera, and markedly so. *Moraxella* interacted with 65 other genera (6.2 SD above the mean), and *Alloiococcus* interacted with 56 other genera (5.2 SD above the mean). *Moraxella* and *Alloiococcus* were, therefore, considered hub genera. Associations with these 2 genera were uniformly negative except for one. *Moraxella* was a node in 40% of all negative associations in the nasal network, and *Alloiococcus* was a node in 35% of all negative associations in the nasal network (Figure 4A and Supplemental Table 6). With 23 associations, *Corynebacterium* also demonstrated connectivity above the mean by 1.8 SD (Figure 4A and Supplemental Table 6). Figure 4B provides more

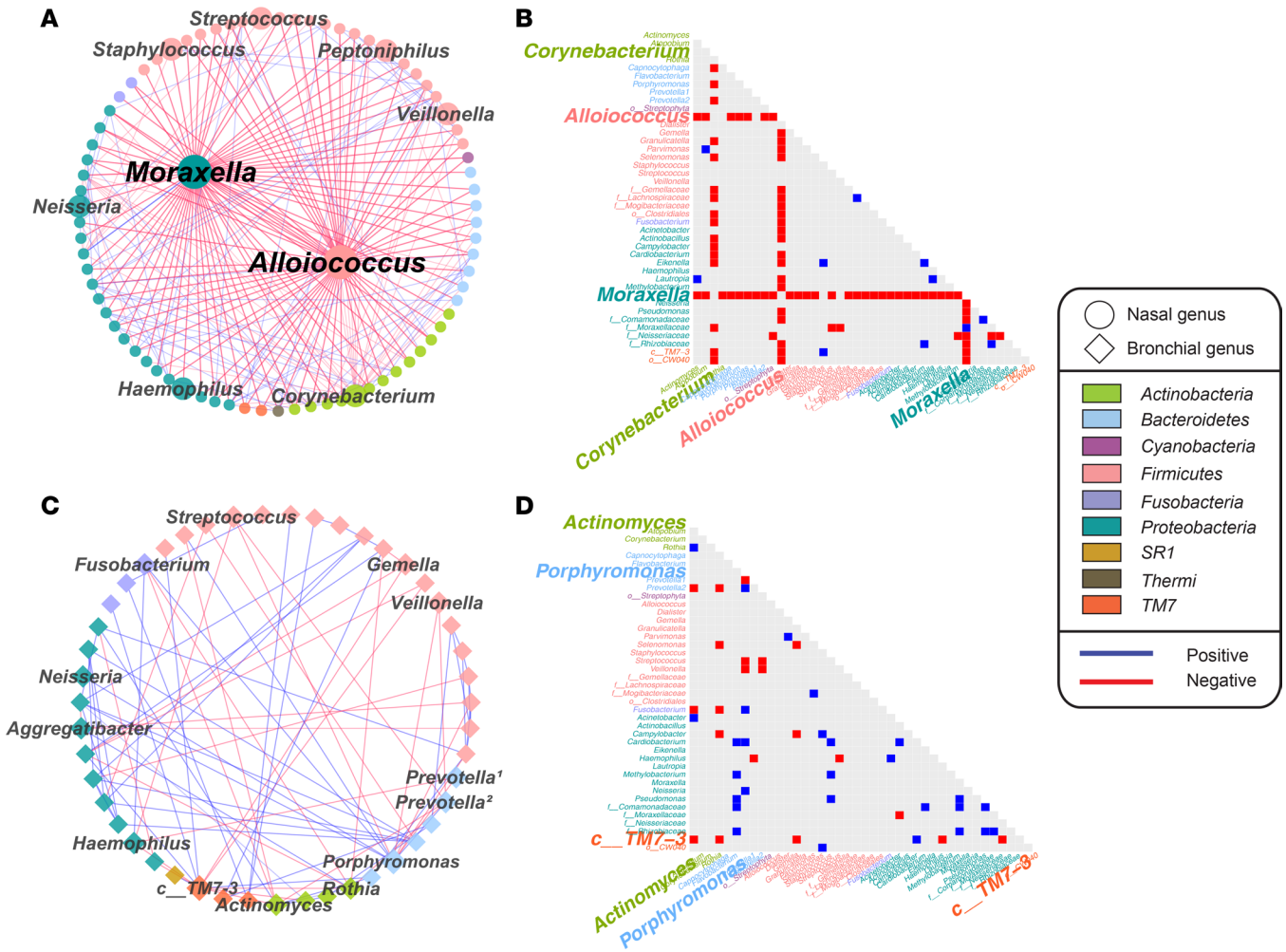


Figure 4. Associations among nasal microbiota and among bronchial microbiota in children with severe persistent asthma (step 2 in Figure 1). (**A**) Network of associations among nasal microbiota in children with severe persistent asthma (n = 27). The network was built based on the ensemble4 method (see Methods). Each circular node represents a nasal genus and is colored by the phylum to which it belongs. Edges between nodes indicate significant associations between genera and are colored based on positive or negative associations between genera abundances. (**B**) Matrix of pairwise associations (FDR ≤ 0.05) between nasal microbiota in children with severe persistent asthma (n = 27). The top 3 most associated genera are shown in bigger font. (**C**) Network of associations between bronchial microbiota in children with severe persistent asthma (n = 27). The network was built using the ensemble4 method (see Methods). Each rhomboid node represents a bronchial genus and is colored by the phyla to which it belongs. Edges between nodes indicate significant associations (FDR ≤ 0.05) between genera and are colored based on positive or negative associations between genera abundances. [1]*Prevotella* genus is affiliated with *Prevotellaceae* family. [2]*Prevotella* genus proposed as affiliated with *Paraprevotellaceae* by the Greengenes database v.13.8. (**D**) Matrix of pairwise associations between bronchial microbiota in children with severe persistent asthma (n = 27). The top 3 most associated genera are shown in bigger font.

detail on the associations between the most connected genera in this nasal network. The matrix highlights that the majority of associations in the nasal network involved *Moraxella*, *Alloiococcus*, or *Corynebacterium*.

The network of associations among microbial genera in the bronchial airway of children with severe persistent asthma is shown in Figure 4C, and the list of all associations in the network is shown in Supplemental Table 7. The mean number of associations for each genus in the bronchial network was 3.8, representing 1.6-fold less mean connectivity per genus relative to the nasal network. In contrast to the nasal network, there were no clear hub genera (Figure 4C). Although an unknown genus in class *TM7-3* and genus *Porphyromonas* had higher numbers of associations than the rest (3.4 and 2.6 SD above the mean, respectively), those numbers did not greatly exceed the average, as was seen for hub genera in the nasal network. For ease of comparison with results from the nasal airway (Figure 4B), Figure 4D provides a matrix of associations between the same genera but in the bronchial airway, with genera most connected in the bronchial airway highlighted with large-font labels. The matrix supports that associations between bronchial genera were not dominated by specific genera, as in the nasal network.

Next, we constructed a network of associations between nasal and bronchial microbiota of children with severe persistent asthma (Figure 5A), and the list of all associations in the network is shown in Supplemental Table 8. We found that bronchial *Porphyromonas* had the highest number of associations with nasal genera (11% of the total number of associations), and nasal *Corynebacterium* had the highest number of associations with bronchial genera (8% of the total number of associations). Nasal-bronchial associations with these hub genera are shown in Figure 5B. We also found that *Cardiobacterium*, *Corynebacterium*, *Eikenella*, *Flavobacterium*, *Haemophilus*, *Lautropia*, an unknown genus in the *Gemellaceae* family, and an unknown genus in class *TM7-3* were each positively coabundant between the nasal and bronchial airways (Figure 5A).

Microbiome-transcriptome networks in the nasal and bronchial airways of severe persistent asthmatics show negative associations between bronchial Actinomyces and bronchial inflammation. To characterize relationships between nasal microbiota and the nasal transcriptome in children with severe persistent asthma (step 3 in Figure 1), we next constructed networks using our nasal microbiome and nasal transcriptome data and an ensemble of 3 statistics (ensemble3 method) derived from permutation-renormalization and bootstrapping (23) of 2 correlation measures (Pearson, Spearman) and the Bray-Curtis dissimilarity measure (24) (see Methods). Summary characteristics of the network constructed are shown in Supplemental Table 9.

The network of associations between nasal microbiota and nasal transcriptome in children with severe persistent asthma is shown in Figure 6A, and the list of all associations in the network is shown in Supplemental Table 10. The mean number of nasally expressed genes associated with each nasal genus in this network was 14.3. *Corynebacterium* interacted with many expressed nasal genes (i.e., 293 associations; 7.8 SD above the mean), representing 30% of the total number of the associations in the nasal microbiome-transcriptome network. Interestingly, all microbiome-transcriptome associations with *Corynebacterium* were negative (i.e., higher *Corynebacterium* abundance associated with lower gene expression of all 293 associated genes). GO analysis of genes associated with *Corynebacterium* revealed no significant enrichments for any biological processes.

Next, to characterize associations between bronchial microbiota and the bronchial transcriptome in children with severe persistent asthma (step 4 in Figure 1), we analogously constructed association networks using our bronchial microbiome and bronchial transcriptome data and the ensemble3 method (Figure 6B). Summary characteristics of this network are shown in Supplemental Table 9, and the list of all associations in the network is shown in Supplemental Table 11. In this bronchial microbiome-transcriptome network, the mean number of genes associated with each bronchial genus was 16.8. *Actinomyces* was the clear hub genus, associated with 157 bronchial genes (5.71 SD above the mean). GO analysis of the bronchial genes associated with *Actinomyces* revealed enrichment for inflammatory response (fold enrichment = 5.7, FDR = 4.9×10^{-4}) (21, 22). Given that all 157 bronchial gene associations with *Actinomyces* were negative, the results from this bronchial network suggest that *Actinomyces* may protect against bronchial inflammation.

Ciliary function genes are overexpressed in nasal airway of children with severe persistent asthma versus healthy controls. We next compared nasal transcriptomes between children with severe persistent asthma and healthy controls (step 5 in Figure 1). A heatmap of the 352 DE genes with FDR ≤ 0.05, and $|\log_2$ fold change$|$ >1 is shown in Figure 7A. GO enrichment analysis revealed that the 266 genes overexpressed in the nasal airways of children with severe persistent asthma versus healthy controls were enriched for processes related to ciliary function, including cilium movement (fold enrichment = 31.1, FDR = 6.9×10^{-6}), outer dynein arm assembly (fold enrichment = 42.0, FDR = 3.0×10^{-4}), inner dynein arm assembly (fold enrichment = 40.4, FDR = 7.9×10^{-3}), and motile cilium assembly (fold enrichment = 35.0, FDR = 1.5×10^{-2}). The 86 genes underexpressed in children with severe persistent asthma versus controls were significantly enriched for epidermis development (fold enrichment = 31.6, FDR = 1.3×10^{-10}), keratinocyte differentiation (fold enrichment = 32.4, FDR = 1.6×10^{-9}), and innate immune response (fold enrichment = 5.7, FDR = 2.6×10^{-2}). The complete list of DE genes in the nasal airway of children with severe persistent asthma versus healthy controls is provided in Supplemental Table 12.

Streptococcus is enriched in the nasal airway of children with severe persistent asthma versus healthy controls. Comparison of the nasal microbiome of children with severe persistent asthma versus healthy controls (step 6 in Figure 1) revealed modest differences. The relative abundances of nasal microbial genera in these 2 groups are shown in Figure 7B. *Corynebacterium* (mean = 30.4%, IQR = 24.3% in healthy; mean = 21.6%, IQR = 25.2% in asthma), *Staphylococcus* (mean = 13.9%, IQR = 14.1% in healthy; mean = 20.9%, IQR = 31.4% in asthma), and *Moraxella* (mean = 23.0%, IQR = 57.3% in healthy; mean = 20.8%, IQR = 23.2% in asthma) were the 3

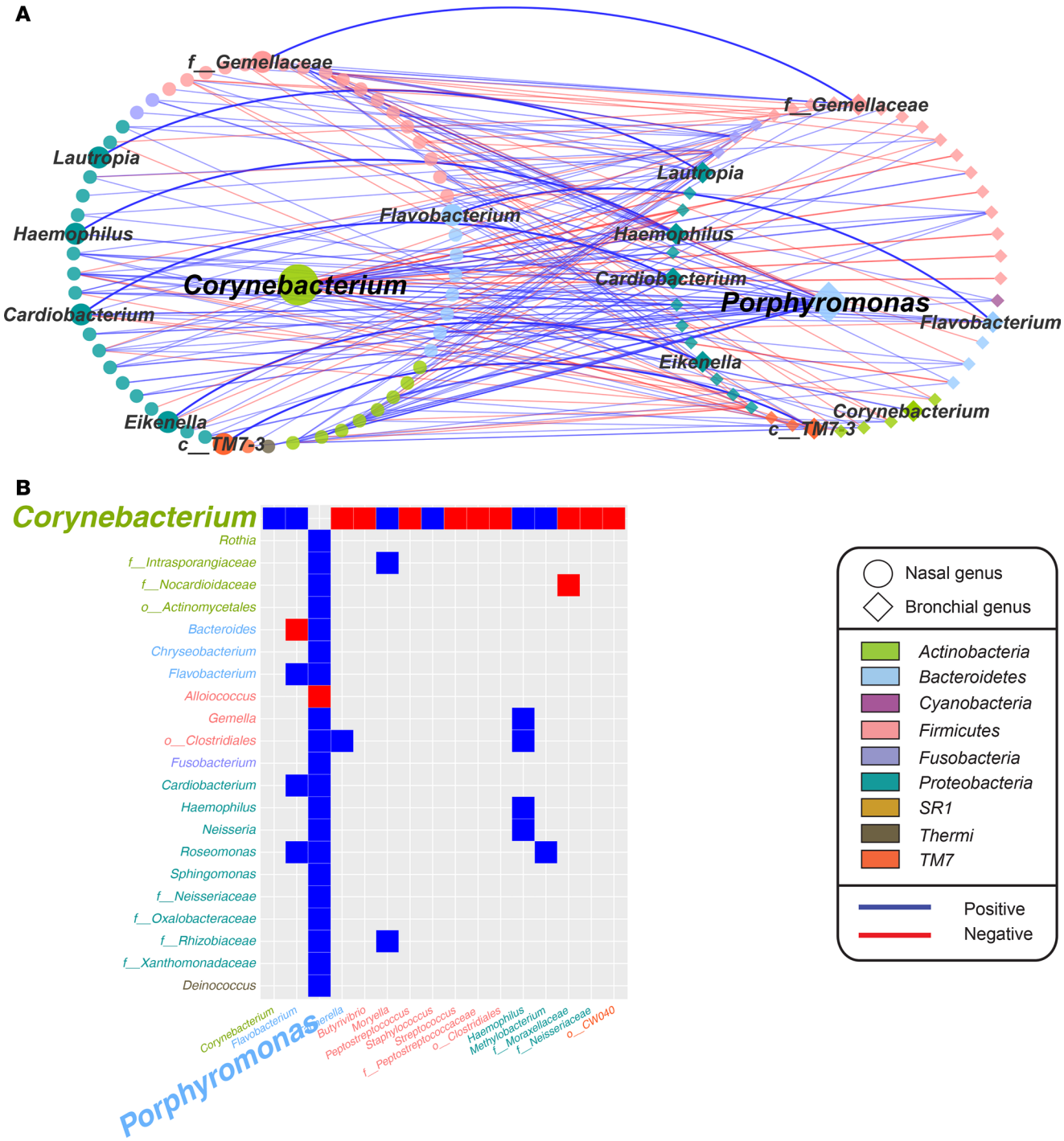


Figure 5. Associations between nasal and bronchial microbiota in children with severe persistent asthma (step 2 in Figure 1). (**A**) Network of associations between nasal microbiota (circular nodes) and bronchial microbiota (rhomboid nodes) in children with severe persistent asthma (n = 27). The network was built using the ensemble4 method (see Methods). Nodes are colored by the phyla to which they belong. Edges indicate significant associations (FDR ≤ 0.05) between genera and are colored based on positive and negative associations between genera abundances. (**B**) Matrix of pairwise associations between nasal and bronchial microbial hubs and their associated genera in children with severe persistent asthma (n = 27). Nasal genera are along the rows, and bronchial genera are along the columns.

most abundant microbial genera in the nasal airways of children with severe persistent asthma and healthy controls. *Streptococcus* was significantly more abundant in the nasal airway of children with severe persistent asthma than in that of healthy controls (LEfSe; ref. 26) (logarithmic Linear Discriminant Analysis [LDA] score = 4.86, $P = 3.26 \times 10^{-2}$). However, there were no significant differences with respect to the following global measures of diversity: α diversity (Shannon diversity index, $P = 0.26$) (Supplemental Figure 2),

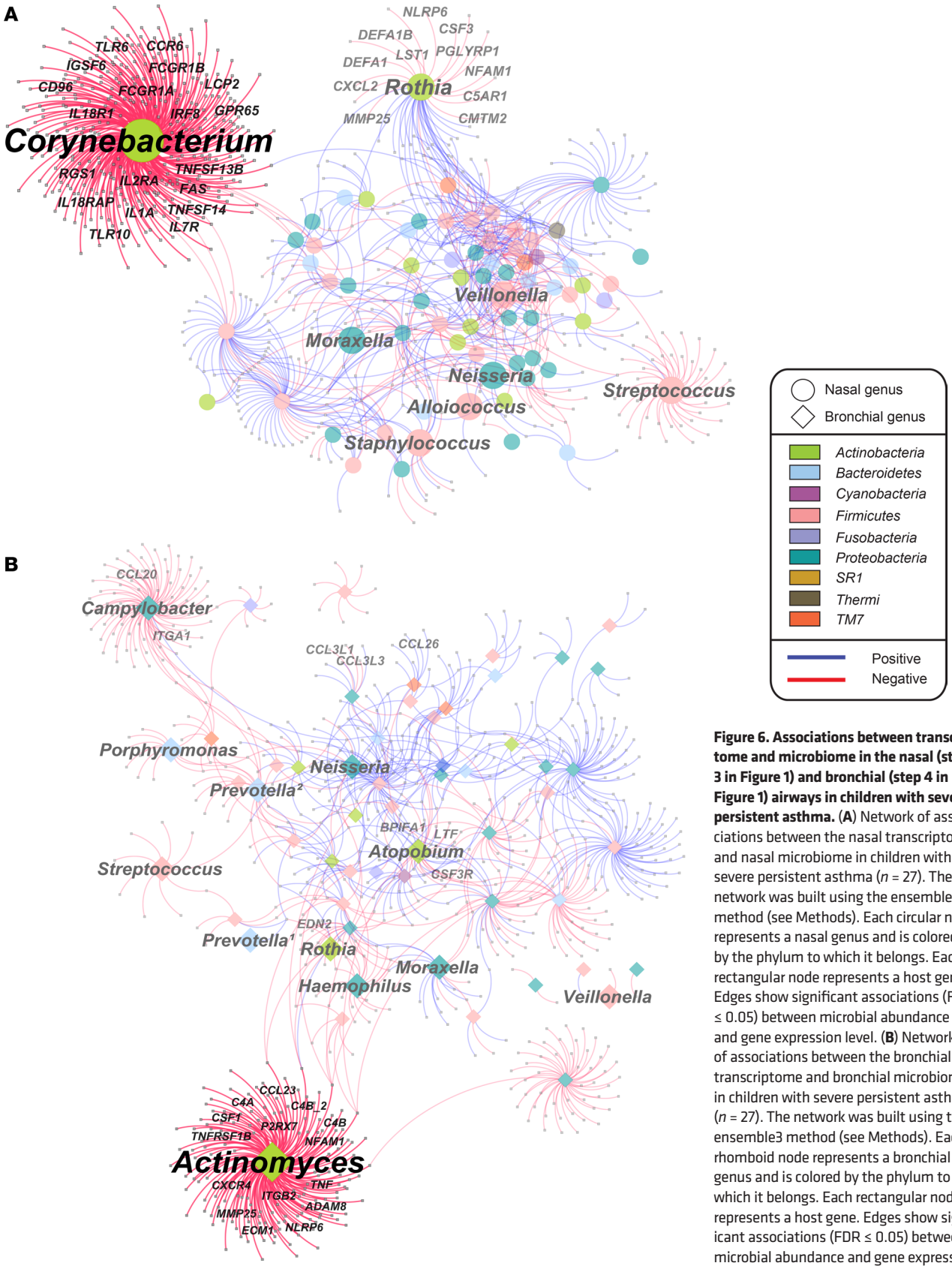


Figure 6. Associations between transcriptome and microbiome in the nasal (step 3 in Figure 1) and bronchial (step 4 in Figure 1) airways in children with severe persistent asthma. (**A**) Network of associations between the nasal transcriptome and nasal microbiome in children with severe persistent asthma (n = 27). The network was built using the ensemble3 method (see Methods). Each circular node represents a nasal genus and is colored by the phylum to which it belongs. Each rectangular node represents a host gene. Edges show significant associations (FDR ≤ 0.05) between microbial abundance and gene expression level. (**B**) Network of associations between the bronchial transcriptome and bronchial microbiome in children with severe persistent asthma (n = 27). The network was built using the ensemble3 method (see Methods). Each rhomboid node represents a bronchial genus and is colored by the phylum to which it belongs. Each rectangular node represents a host gene. Edges show significant associations (FDR ≤ 0.05) between microbial abundance and gene expression level. [1]*Prevotella* genus is affiliated with *Prevotellaceae* family. [2]*Prevotella* genus proposed as affiliated with *Paraprevotellaceae* by Greengenes database v.13.8.

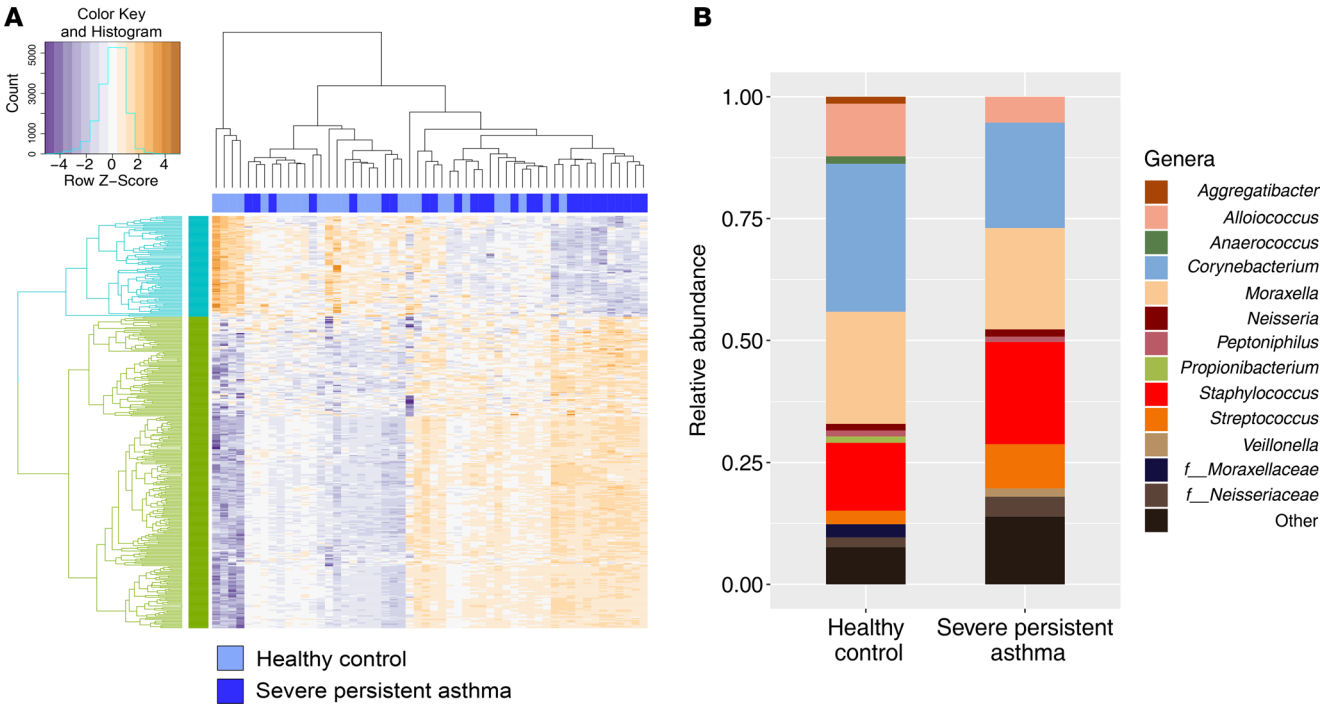


Figure 7. Nasal transcriptome (step 5 in Figure 1) and microbial composition (step 6 in Figure 1) of healthy children versus children with severe persistent asthma. (**A**) Heatmap of nasal genes differentially expressed (FDR ≤ 0.05, and $|\log_2$ fold change$|$ > 1) between healthy children (n = 27) and children with severe persistent asthma (n = 27). Rows represent 86 genes upregulated in healthy controls versus children with severe persistent asthma (cyan), and 266 genes upregulated in children with severe persistent asthma versus healthy controls (green). Columns indicate the 27 healthy control (light blue) and 27 severe persistent asthma (dark blue) nasal samples. (**B**) Relative abundance of nasal bacterial genera of healthy children (n = 27) versus children with severe persistent asthma (n = 27). Genera representing < 1% of total abundance are shown as Other.

β diversity (unweighted UniFrac; 1-way PERMANOVA P = 0.21) (Supplemental Figure 3A), and weighted UniFrac (1-way PERMANOVA P = 0.17) (Supplemental Figure 3B).

Comparison of asthma nasal microbiome network and healthy nasal microbiome network shows similarity between hub genera but distinct associations. To characterize relationships between microbiota within the nasal airway of healthy children, we next built a network using nasal microbiome data from our healthy controls (Figure 8A). The characteristics of this healthy nasal microbiome network are shown in Supplemental Table 5, and the list of all associations in the network is shown in Supplemental Table 13.

We compared this nasal microbiome network from healthy children (Figure 8A) with the nasal microbiome network from children with severe persistent asthma (Figure 4A). As was the case in the asthma nasal microbiome network (Figure 4A), *Moraxella* and *Alloiococcus* were hub genera in the healthy nasal microbiome network (Figure 8A) with 58 associations (6.2 SD above the mean) and 42 associations (4.3 SD above the mean), respectively. Similarly, *Corynebacterium* was the third most associated genus, with 27 associations (2.5 SD above the mean). Figure 8B provides a matrix of the associations among genera in this healthy nasal microbiome network.

Although hub genera appeared similar, only 95 microbial associations (41% of asthma nasal microbiome network associations and 36% of healthy nasal microbiome network associations) overlapped between the asthma and healthy nasal microbiome networks. Interestingly, associations with the top 3 most associative genera (*Moraxella*, *Alloiococcus*, *Corynebacterium*) were much better conserved between the 2 networks than associations with other genera. We recognize that differential connectivity may not necessarily indicate differential functionality.

Comparison of nasal microbiome-transcriptome networks from healthy and asthmatic children reveals that Corynebacterium is negatively associated with inflammation-related biologic processes in healthy but not asthmatic children. As a final step to our systematic study, we constructed a healthy nasal microbiome-transcriptome network (Figure 9) using nasal microbiome and nasal host transcriptome profiles from our healthy controls (step 7 in Figure 1). The characteristics of this network are summarized in Supplemental Table 9, and the list of all

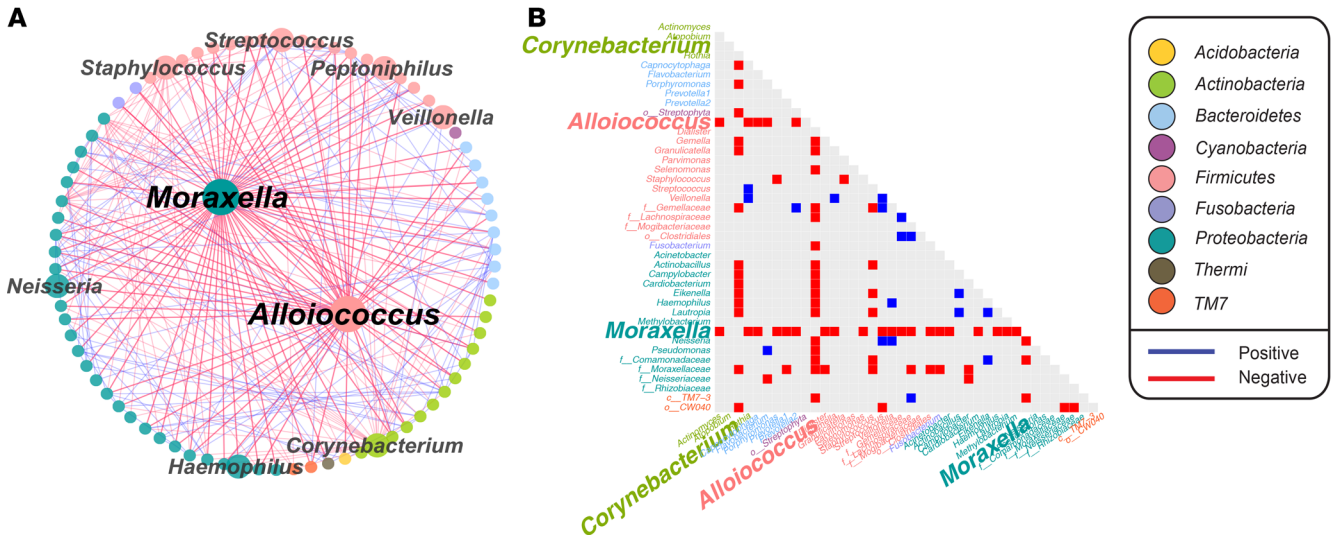


Figure 8. Associations between nasal microbiota in healthy controls. (**A**) Nasal microbial network in healthy children (n = 27). The network was built using the ensemble4 method (see Methods). Each circular node represents a nasal genus and is colored by the phylum to which it belongs. Edges between nodes indicate significant associations (FDR ≤ 0.05) between genera and are colored based on positive and negative associations between genera abundances. (**B**) Matrix of pairwise associations between nasal microbiota in healthy children (n = 27). The top 3 most associated genera are shown in bigger font.

associations is shown in Supplemental Table 14. The mean number of genes associated with each genus in this healthy nasal microbiome-transcriptome network was 21.9. As in the analogous network for children with severe persistent asthma (Figure 6A), *Corynebacterium* demonstrated an exceptionally large number of associations with the host nasal genes, with 565 associations (7.9 SD above the mean). All 565 associations with *Corynebacterium* were negative except for 2, and the number of associations with *Corynebacterium* was 48% of the total number of negative associations in the network.

In contrast to the asthma nasal microbiome-transcriptome network where member genes were not enriched for any biological process, genes associated with *Corynebacterium* in this healthy nasal microbiome-transcriptome network were significantly enriched for a large number of immune and inflammation-related GO biological processes. The top GO terms based on analyses using DAVID (21, 22), ranked by FDR, were inflammatory response (fold enrichment = 5.7, FDR = 1.8×10^{-24}), immune response (fold enrichment = 4.8, FDR = 7.2×10^{-19}), and chemotaxis (fold enrichment = 7.7, FDR = 6.0×10^{-12}). This suggests that, in the healthy nasal airway, *Corynebacterium* may be more protective against inflammation, since it negatively interacts with genes that promote inflammation. The fact that such potentially beneficial suppressive associations were observed and found to be statistically significant as a set in the healthy network but not in the asthma network suggests that such potential protection may be impaired in children with severe persistent asthma. Further comparison of the healthy and asthma nasal microbiome-transcriptome networks revealed that, although their summary characteristics appeared similar (Supplemental Table 9), only 26 among more than a thousand edges overlapped between the 2 networks. Additionally, the number of associations with host genes per microbial genus was much higher in the healthy network (mean 21.9, SD = 69.1) than in the asthma network (mean 14.3, SD = 35.7), indicating that microbiota in the healthy nasal airway have higher connectivity with host gene expression.

Discussion

While host and microbial factors are each known to influence asthma (5, 27), deciphering their interactions has been more challenging. This study's integrated study of parallel transcriptome and microbiome profiles from the nasal and bronchial airways of children with severe persistent asthma and healthy controls has enabled progress in understanding relationships between the airway microbiome and airway host transcriptome in asthma. While there have been studies of the airway microbiome in asthma (5–14), these have typically taken place separately from studies of the airway transcriptome in asthma (5, 15–18).

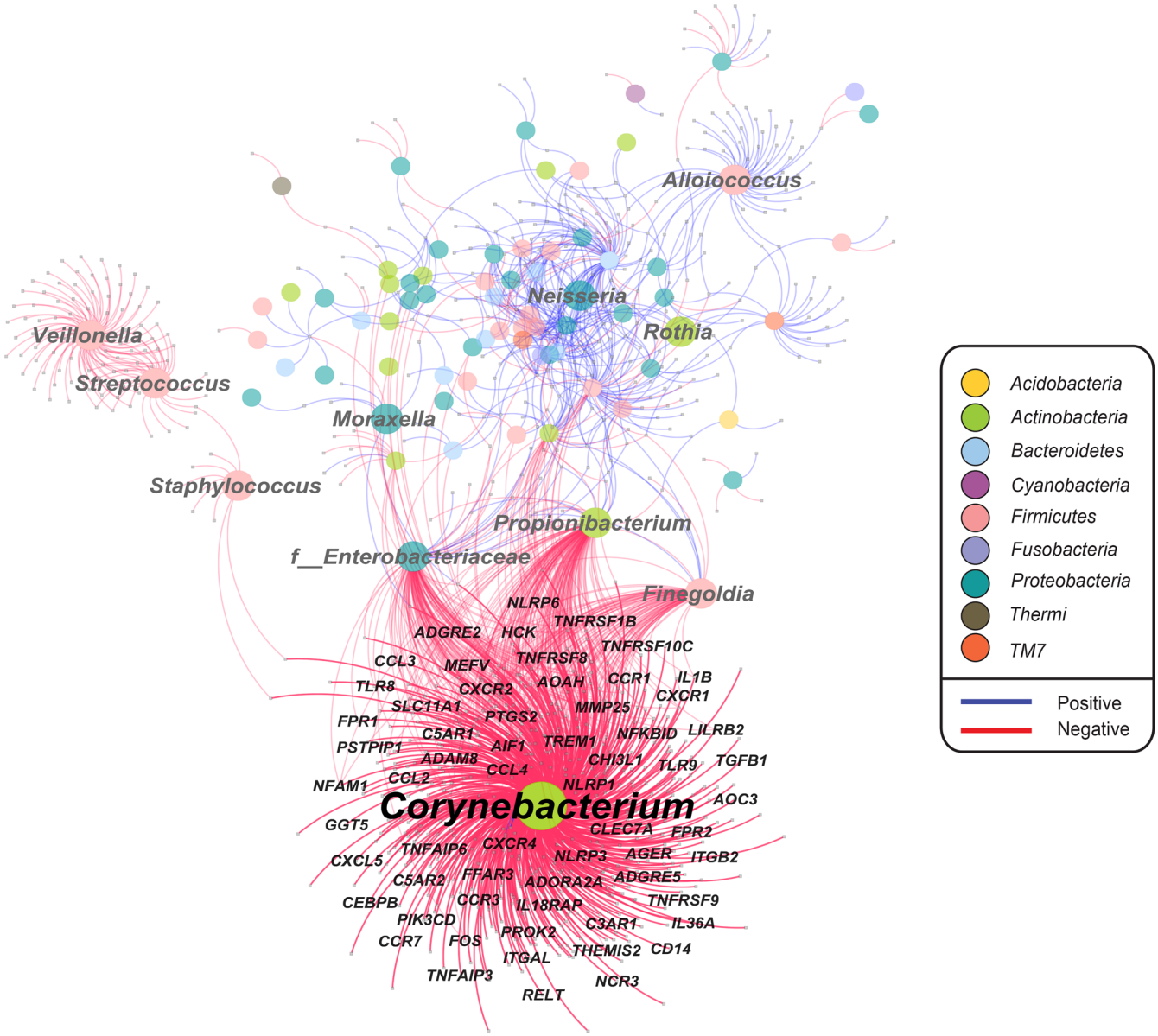


Figure 9. The nasal transcriptome and nasal microbiome in healthy controls (step 7 in Figure 1). Network of associations between the nasal transcriptome and nasal microbiome in healthy children (n = 27). The network was built using the ensemble3 method (see Methods). Each circular node represents a nasal genus and is colored by the phylum to which it belongs. Each rectangular node represents a host gene. Edges show significant associations (FDR ≤ 0.05) between microbial abundance and gene expression level.

Additionally, whether from the microbiome lens or the transcriptome lens, the nasal and bronchial airways have largely been examined separately, even though asthma affects and is affected by the entire airway. Our systematic, parallel profiling of the nasal and bronchial microbiomes and transcriptomes of children with asthma allowed for unprecedented integrated analysis of relationships between microbiota and host transcriptome in the nasal and bronchial airways of the same children. These relationships would have been be detected by microbiome-only or transcriptome-only studies, or by nasal-only or bronchial-only studies. Additionally, our study design enabled the uncovering of results about the asthmatic airway unbiased by interindividual differences.

We found that *Moraxella* and *Alloiococcus* were hub genera in the nasal microbiome networks of children with severe persistent asthma (Figure 4A) and healthy controls (Figure 8A), while the bronchial microbiome network from our study (Figure 4C) demonstrated no hubs. Previous work (28, 29) supports that microbial communities with at least 1 hub microbe are less susceptible to perturbation by host or

environmental factors. Interestingly, we found 2 hub nodes in each of our nasal microbiome networks, suggesting that these nasal microbial communities may be less susceptible to disturbance than the bronchial microbial community, where the respective network did not reveal any hub genera. This may speak to the fact that the nasal airway is the first point of entry to the respiratory tract and may therefore benefit from being resilient to perturbation. Although the bronchial microbiome network did not follow a scale-free structure typically observed with biological networks, recent studies support that many biological networks are often not scale free (30), including microbial networks (31, 32).

Moraxella, a nasal genus that has been previously associated with asthma exacerbation (7) and the development of asthma (15), was the most associated genus in the nasal airways of children with severe persistent asthma (Figure 4, A and B). The second and the third most associated genera were *Alloiococcus* and *Corynebacterium*, which have been previously associated with respiratory health (6, 33). Interestingly, our systematic analysis showed similar hub genera in the nasal microbiome of healthy children (Figure 8A). Although hub genera of the nasal microbiome networks for asthmatics and for healthy controls appeared similar, the majority of associations with hub genera were distinct between the asthma and healthy nasal microbiome networks, supporting that the dynamics of associations between genera may better distinguish disease state than just the presence and/or abundance of a genus. Notwithstanding, we did find that *Streptococcus* was enriched in the nasal airway of children with severe persistent asthma versus controls, consistent with prior observations that its abundance is associated with asthma exacerbation (7) and asthma development (15) in children. Additionally, studies of farm environments and asthma risk have shown that the abundance of *Streptococcaceae* in dust microbiota is higher in nonfarm homes associated with asthma risk (34).

In the nasal microbiome-transcriptome network of healthy children (Figure 9), *Corynebacterium* had numerous negative associations with a large number of genes related to inflammation, suggesting that the genus could be protective against airway inflammation. This would be consistent with prior work demonstrating that *Corynebacterium* in the upper airway is negatively associated with proinflammatory cytokines in BAL (13) and positively associated with the elimination of potential pathogens and respiratory health (33). In the nasal microbiome-transcriptome network of children with severe persistent asthma (Figure 6A), *Corynebacterium* also interacted with genes that individually play a role in inflammation, but the number of these genes was far smaller than in the healthy network, and together as a set, these *Corynebacterium*-associated genes were not significantly enriched for inflammatory processes. This suggests that, relative to healthy children, the protective effects of *Corynebacterium* are attenuated in children with severe persistent asthma. Our findings support that the *Corynebacterium* species is a likely beneficial commensal bacterium that is abundant in the nasal airway and is associated with suppression of proinflammatory genes.

In the bronchial microbiome-transcriptome network of children with severe persistent asthma (Figure 6B), *Actinomyces* (a genus in the phylum *Actinobacteria* that also includes *Corynebacterium*) was a hub genus with negative associations with genes for inflammatory response. Thus, similar to *Corynebacterium* in the nasal airway, *Actinomyces* in the bronchial airway may be protective against bronchial inflammation. As we were not ethically able to perform bronchoscopy on healthy children, we cannot say how these associations would compare (e.g., be fewer) to the associations that would be seen for *Actinomyces* if a healthy bronchial microbiome-transcriptome network could be constructed. Consistent with our interpretation of *Actinomyces* as antiinflammatory bronchial microbiota, however, a study of adults with mild atopic asthma found that the relative abundance of *Actinomyces* in bronchial brushing samples was negatively associated with serum IgE levels and BAL neutrophil counts (13).

A limitation to this study was that the number of subjects examined may have limited our power to detect additional findings. However, our sample exceeds the size of previous studies of the airway microbiome in children where anesthesia was required to obtain biospecimens (8, 9, 20). Furthermore, we systematically studied each child with severe persistent asthma by 4 molecular dimensions (nasal microbiome, bronchial microbiome, nasal transcriptome, and bronchial transcriptome), thus moving beyond the scope of what prior studies have attempted. Also, while it would have been scientifically compelling to obtain bronchial samples from healthy children to further extend our systematic study, this was not ethically possible due to the fact that bronchoscopy is an invasive procedure that is not indicated in healthy children. Last, while our study represents an important step forward in systematically characterizing associations and correlations between upper and lower airway microbiota and host gene expression in children with asthma, the causal directions of these relationships could not be established with our study design.

Our systematic, integrative study provides a window into the nasal and bronchial microbiome, nasal and bronchial transcriptome, and their associations, in children with severe persistent asthma and healthy controls.

Future work with larger sample sizes could enable analyses for causal interactions between airway microbiota and host transcriptomes, and additional work incorporating viral and fungal profiling could expand systematic studies such as this and refine our understanding of host-microbe interactions in asthma.

Methods

Subject recruitment and sample collection. Children with severe persistent asthma were recruited from the pediatric pulmonology practice within the Mount Sinai Health System (New York, New York, USA). Children were determined to have severe persistent asthma based on criteria proposed by an American Thoracic Society (ATS) workshop on refractory asthma (35) that were adopted by the Severe Asthma Research Program (SARP) (36). These criteria were centered around symptom intensity, medication requirements, treatment needs, and lung function results (36). The children with severe persistent asthma in our cohort underwent flexible bronchoscopy for clinical management of their refractory asthma.

Healthy controls were recruited from general pediatrics clinics in the Mount Sinai Health System. Any child with allergic rhinitis, or personal or family history of any airway disease, was excluded. All healthy controls were required to demonstrate before and after bronchodilator spirometry within normal limits.

Nasal and bronchial transcriptome profiling. For transcriptome profiling of nasal and bronchial airways of children with severe persistent asthma, and the nasal airway of healthy controls, nasal brushing was performed with a standard cytology brush for all subjects, and bronchial brushing was performed using a bronchoscopy brush for the subjects with asthma. Brushes were immediately placed in RNALater (Thermo Fisher Scientific). RNA extraction was performed with the Qiagen RNeasy Mini Kit. Samples were assessed for yield and quality using the 2100 Bioanalyzer (Agilent Technologies) and Qubit fluorometry (Thermo Fisher Scientific). The RNA-seq library was prepared with the standard TruSeq RNA Sample Prep Kit v2 protocol (Illumina).The mRNA libraries were sequenced on the Illumina HiSeq 2500 platform with a per-sample target of 40 million–50 million 100 bp paired-end reads. The data were put through Mount Sinai's standard mapping pipeline (37) (using Bowtie and TopHat and assembled into gene — and transcription — level summaries using Cufflinks; refs. 38–40). Mapped data were subjected to quality control with tissFastQC and RNA-SeQC (41). R package edgeR (42) was used to normalize the dataset after filtering out genes with low expression by retaining genes with counts per million (CPM) > 1 in at least 1 quarter of the total number of samples.

Nasal and bronchial microbiome profiling. For microbiome profiling, nasal samples were obtained from children with severe persistent asthma and healthy controls using a sterile cotton swab, and bronchial samples were obtained from children with severe persistent asthma by BAL of the right middle lobe during bronchoscopy. As a precaution for potential false positives from contaminants, we included blank swab controls with the clinical collections. DNA extraction from all swabs was performed using Qiagen DNeasy Mini Kit. Sample DNA yield and quality was accessed using Nanodrop (Thermo Fisher Scientific). 16S rRNA library preparation was done according to the protocol described in the Illumina MiSeq 16S Metagenomic Sequencing Library Preparation Protocol. Both agarose DNA gel electrophoresis and Bioanalyzer were used for analysis and quantification of the final PCR product. Dual indices and Illumina sequencing adapters were attached to the V3–V4 amplicon of 16S rRNA using the Nextera XT Index Kit v2 (sets A and B, Illumina). PCR blanks were included with the samples for sequencing as negative controls. DNA sequencing was performed on amplified 16S rRNA using the Illumina MiSeq platform with 2 × 250bp paired-end reads. Primer regions and the low-quality bp at the end of the reads were trimmed using SEQTK 1.2 (43). Reads were demultiplexed and filtered with the minimum acceptable Phred quality score of 20 using Qiime 1.9.1 (44). Operational taxonomic units (OTUs) were picked by clustering sequences at a 97% similarity threshold also using Qiime 1.9.1 (44) with Greengenes reference data version 13.8. We subtracted OTU counts detected in the blank samples to remove signal from potential contaminants.

Transcriptome analyses. The Principal Component Analysis (PCA) plots were generated using R (version 3.5.1, R Foundation for Statistical Computing). DE genes between the nasal and bronchial samples of children with severe persistent asthma were analyzed as paired samples using R packages voom (45) and limma (46) with FDR threshold ≤ 0.05 and $|\log_2$ fold change$|$ > 1. Differential gene expression analysis of nasal transcriptomes between children with severe persistent asthma and healthy controls was similarly performed but as independent samples. Heatmaps for DE genes were plotted using R. The enrichments of the GO terms were analyzed using DAVID 6.8 (21, 22).

Microbiome diversity and composition analyses. Both α and β diversities and genera composition analyses were performed using Qiime 1.9.1 (44). To compare α diversities measured by Shannon diversity index and to compare relative abundances in genera between the nasal and bronchial microbiomes in children with severe persistent asthma, Wilcoxon signed-rank test for paired samples was performed, and *P* values were corrected by FDR. To compare β diversities between groups, unweighted and weighted UniFrac distances were used for 1-way PERMANOVA tests with 1000 permutations. To compare α diversities between the nasal microbiome in children with severe persistent asthma versus healthy controls measured by Shannon diversity index, the Monte Carlo method was used with 1000 permutations, with correction for multiple comparisons by FDR. LEfSe (26) was used to calculate differentially abundant genera between these 2 groups.

Microbiome network construction. Normalized 16S rRNA profiling summarizes bacterial abundance as relative, meaning an increase in relative abundance of one taxon leads to a decrease in that of another (23). Because of this, it is difficult to distinguish whether a negative association measured based on relative abundance data represents a real competition between bacteria or false information due to flourishing of one bacterium while the populations of others remain the same (23). Therefore, we used the permutation-renormalization and bootstrap (ReBoot) method, which constructs a null distribution using permutation-renormalization that reflects the compositional effect in relative abundance data and compares it with a bootstrapped real distribution to suppress compositional effects in bacterial 16S profiling while estimating the significance of associations (23).

Another characteristic of 16S rRNA data is sparsity. Correlation-based analyses are sensitive to the double-zero problem, meaning a zero count may represent either true absence of taxa or a result of undersampling so that comparing 2 zeros is not informative (23, 47, 48). To overcome the limitation of correlation-based analysis on sparse 16S data, we used an ensemble of 4 different measures (ensemble4 method) (23) including 2 measures of correlation (Pearson and Spearman) and 2 measures of dissimilarity (Bray-Curtis dissimilarity and Kullback-Leibler divergence; refs. 24 and 25). We use the word "association" to refer to relationships significant by the above ensemble statistic. Before constructing the networks, genera with zero counts in more than two-thirds of the total number of samples of each airway site were removed (23). Permutations and bootstraps were performed 1000 times each, and then the 2 distributions were compared using a z-test for each of the 4 methods (23). The associations between bacterial genera from each of the 4 methods were ranked by correlation coefficients or dissimilarity values and prefiltered so that, for each measure, the top 20% and bottom 20% of associations were identified for input into the downstream network reconstruction algorithm. This prefiltering step removed high FDR associations. Correlations or dissimilarity measures between genera that survived the prefiltering of all 4 measures were included for further analyses. The *P* values from the 4 measures were combined using Simes method (49), and the combined *P* values were FDR corrected. The network was then defined by genera as nodes and edges between the nodes indicating associations with FDR ≤ 0.05. Associations with FDR > 0.05 or with inconsistent Pearson and Spearmen correlation directions were filtered out. Cytoscape 3.5.1 (50) was used for network visualization.

Microbiome-transcriptome network construction. Genera with zero counts in more than two-thirds of the total number of samples of each airway site were removed (23). We used an ensemble of 3 different measures (ensemble3 method) using the ReBoot method (23) spanning Spearman correlation, Pearson correlation, and Kullback-Leibler divergence (25) analyses to assess associations between the normalized transcriptome data and the microbiome relative abundance data. Compared with the ensemble4 method used for the microbiome network construction, the Bray-Curtis dissimilarity (24) was excluded from ensemble3 because it is designed to quantify the compositional dissimilarity of species populations between 2 sites, whereas transcriptome data are not compositional. The downstream methods for constructing networks of associations between microbiome and their host transcriptome were the same as the methods for the networks of associations between genera, except the prefiltering threshold for each of the 3 methods was made more stringent for the top 3% and bottom 3% to remove high FDR associations. Cytoscape 3.5.1 (50) was used for network visualization.

Data availability. Data for this study are available at the NCBI Sequence Read Archive (SRA) (https://www.ncbi.nlm.nih.gov/sra), accession no. PRJNA601756 for transcriptome data and accession no. PRJNA601757 for microbiome data.

Statistics. The statistics used in this study are detailed in the following sections above: Transcriptome Analyses, Microbiome Diversity and Composition Analyses, Microbiome Network Construction, and Microbiome-transcriptome Network Construction.

Study approval. Written informed consent was obtained from the parents of all participants, and assent was received from all participants. The study was approved by the Mount Sinai IRB.

Author contributions

SB directed, designed, supervised, and procured funding for this study. SB, SR, CS, and AV worked on the recruitment of subjects and sample collection. GG and AG coordinated sample intake, processing, and sequencing. SB, YC, and AD designed and performed the statistical and computational analyses. ES and GF provided guidance on the analyses. SB and YC wrote the original draft of the manuscript. All authors critically reviewed and edited the manuscript. All authors contributed to the work presented in this paper.

Acknowledgments

We thank the children who participated in this study. This study was funded by the US NIH R01 AI 118833.

Address correspondence to: Supinda Bunyavanich, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue #1498, New York, New York 10029, USA. Phone: 212.659.8262; Email: supinda@post.harvard.edu.

1. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2015;386(9995):743–800.
2. Chung KF, et al. International ERS/ATS guidelines on definition, evaluation and treatment of severe asthma. *Eur Respir J*. 2014;43(2):343–373.
3. Fitzpatrick AM. Severe Asthma in Children: Lessons Learned and Future Directions. *J Allergy Clin Immunol Pract*. 2016;4(1):11.
4. Fitzpatrick AM, Moore WC. Severe Asthma Phenotypes - How Should They Guide Evaluation and Treatment? *J Allergy Clin Immunol Pract*. 2017;5(4):901–908.
5. Tyler SR, Bunyavanich S. Leveraging -omics for asthma endotyping. *J Allergy Clin Immunol*. 2019;144(1):13–23.
6. Teo SM, et al. The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development. *Cell Host Microbe*. 2015;17(5):704–715.
7. McCauley K, et al. Distinct nasal airway bacterial microbiotas differentially relate to exacerbation in pediatric patients with asthma. *J Allergy Clin Immunol*. 2019;144(5):1187–1197.
8. Hilty M, et al. Disordered microbial communities in asthmatic airways. *PLoS One*. 2010;5(1):e8578.
9. Kloepfer KM, et al. In children, the microbiota of the nasopharynx and bronchoalveolar lavage fluid are both similar and different. *Pediatr Pulmonol*. 2018;53(4):475–482.
10. Huang YJ, et al. The airway microbiome in patients with severe asthma: Associations with disease features and severity. *J Allergy Clin Immunol*. 2015;136(4):874–884.
11. Fazlollahi M, et al. The nasal microbiome in asthma. *J Allergy Clin Immunol*. 2018;142(3):834–843.e2.
12. Durack J, et al. Features of the bronchial bacterial microbiome associated with atopy, asthma, and responsiveness to inhaled corticosteroid treatment. *J Allergy Clin Immunol*. 2017;140(1):63–75.
13. Durack J, et al. Bacterial biogeography of adult airways in atopic asthma. *Microbiome*. 2018;6(1):104.
14. Stokholm J, et al. Maturation of the gut microbiome and risk of asthma in childhood. *Nat Commun*. 2018;9(1):141.
15. Bisgaard H, et al. Childhood asthma after bacterial colonization of the airway in neonates. *N Engl J Med*. 2007;357(15):1487–1495.
16. Pandey G, et al. A Nasal Brush-based Classifier of Asthma Identified by Machine Learning Analysis of Nasal RNA Sequence Data. *Sci Rep*. 2018;8(1):8826.
17. Poole A, et al. Dissecting childhood asthma with nasal transcriptomics distinguishes subphenotypes of disease. *J Allergy Clin Immunol*. 2014;133(3):670–8.e12.
18. Modena BD, et al. Gene Expression Correlated with Severe Asthma Characteristics Reveals Heterogeneous Mechanisms of Severe Disease. *Am J Respir Crit Care Med*. 2017;195(11):1449–1463.
19. Yii ACA, Tay TR, Choo XN, Koh MSY, Tee AKH, Wang DY. Precision medicine in united airways disease: A "treatable traits" approach. *Allergy*. 2018;73(10):1964–1978.
20. An SQ, Warris A, Turner S. Microbiome characteristics of induced sputum compared to bronchial fluid and upper airway samples. *Pediatr Pulmonol*. 2018;53(7):921–928.
21. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44–57.
22. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009;37(1):1–13.
23. Faust K, et al. Microbial co-occurrence relationships in the human microbiome. *PLoS Comput Biol*. 2012;8(7):e1002606.
24. Bray JR, Curtis JT. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecol Monogr*. 1957;27(4):325–347.
25. Kullback S, Leibler RA. On Information and Sufficiency. *Ann Math Stat*. 1951;22(1):79–86.
26. Segata N, et al. Metagenomic biomarker discovery and explanation. *Genome Biol*. 2011;12(6):R60.
27. Noval Rivas M, Crother TR, Arditi M. The microbiome in asthma. *Curr Opin Pediatr*. 2016;28(6):764–771.
28. Agler MT, et al. Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. *PLoS Biol*. 2016;14(1):e1002352.

29. Hu JX, Thomas CE, Brunak S. Network biology concepts in complex disease comorbidities. *Nat Rev Genet*. 2016;17(10):615–629.
30. Broido AD, Clauset A. Scale-free networks are rare. *Nat Commun*. 2019;10(1):1017.
31. Parente E, Zotta T, Faust K, De Filippis F, Ercolini D. Structure of association networks in food bacterial communities. *Food Microbiol*. 2018;73:49–60.
32. Sam Ma Z, Guan Q, Ye C, Zhang C, Foster JA, Forney LJ. Network analysis suggests a potentially 'evil' alliance of opportunistic pathogens inhibited by a cooperative network in human milk bacterial communities. *Sci Rep*. 2015;5:8275.
33. Man WH, de Steenhuijsen Piters WA, Bogaert D. The microbiota of the respiratory tract: gatekeeper to respiratory health. *Nat Rev Microbiol*. 2017;15(5):259–270.
34. Kirjavainen PV, et al. Farm-like indoor microbiota in non-farm homes protects children from asthma development. *Nat Med*. 2019;25(7):1089–1095.
35. [No authors listed]. Proceedings of the ATS workshop on refractory asthma: current understanding, recommendations, and unanswered questions. American Thoracic Society. *Am J Respir Crit Care Med*. 2000;162(6):2341–2351.
36. Moore WC, et al. Characterization of the severe asthma phenotype by the National Heart, Lung, and Blood Institute's Severe Asthma Research Program. *J Allergy Clin Immunol*. 2007;119(2):405–413.
37. Fromer M, et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat Neurosci*. 2016;19(11):1442–1453.
38. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25.
39. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009;25(9):1105–1111.
40. Trapnell C, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28(5):511–515.
41. DeLuca DS, et al. RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics*. 2012;28(11):1530–1532.
42. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139–140.
43. Li H. Toolkit for processing sequences in FASTA/Q formats. GitHub. https://github.com/lh3/seqtk. Accessed February 19, 2020.
44. Caporaso JG, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010;7(5):335–336.
45. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15(2):R29.
46. Ritchie ME, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
47. Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. *Nat Methods*. 2013;10(12):1200.
48. Legendre P, Legendre LF. *Numerical Ecology*. Oxford, UK; Elsevier; 2012.
49. Simes RJ. An Improved Bonferroni Procedure for Multiple Tests of Significance. *Biometrika*. 1986;73(3):751–754.
50. Shannon P, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–2504.