Respiratory microbiome profiles are associated with distinct inflammatory phenotype and lung function in children with asthma

Running Title: Lung microbiome for asthmatic children

Kim YH\textsuperscript{1,2}, Park MR\textsuperscript{1,2}, Kim SY\textsuperscript{2,3}, Kim MY\textsuperscript{2,4}, Kim KW\textsuperscript{2,3}, Sohn MH\textsuperscript{2,3}

\textsuperscript{1}Department of Pediatrics, Gangnam Severance Hospital, Seoul
\textsuperscript{2}Institute of Allergy, Severance Biomedical Science Institute, Brain Korea 21 Project for Medical Science, Yonsei University College of Medicine, Seoul
\textsuperscript{3}Department of Pediatrics, Severance Hospital, Seoul
\textsuperscript{4}Department of Pediatrics, Yongin Severance Hospital, Yongin, Korea

Corresponding author:
Myung Hyun Sohn
Department of Pediatrics, Severance Hospital, Institute of Allergy, Brain Korea 21 PLUS Project for Medical Science, Yonsei University
College of Medicine, Seoul, Korea
E-mail: mhsohn@yuhs.ac

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.18176/jiaci.0918
ABSTRACT

Background: Respiratory microbiome studies have fostered our understanding of various phenotypes and endotypes of heterogeneous asthma. However, the relationship between the respiratory microbiome and clinical phenotypes in children with asthma remains unclear. We aimed to identify microbiome-driven clusters reflecting the clinical features of asthma and their dominant microbiotas in children with asthma.

Methods: Induced sputum was collected from children with asthma, and microbiome profiles were generated via sequencing of the V3–V4 region of the 16S rRNA gene. Cluster analysis was performed using the partitioning around medoid clustering method. The dominant microbiota in each cluster was determined using the Linear Discriminant Effect Size analysis. Each cluster was analyzed for association among the dominant microbiota, clinical phenotype, and inflammatory cytokine.

Results: Eighty-three children diagnosed with asthma were evaluated. Among four clusters reflecting the clinical characteristics of asthma, cluster 1, dominated by Haemophilus and Neisseria, demonstrated lower post-bronchodilator (BD) forced expiratory volume in 1 second (FEV1)/forced vital capacity (FVC) than that in the other clusters and more mixed granulocytic asthma. Neisseria negatively correlated with pre-BD and post-BD FEV1/FVC. Haemophilus and Neisseria positively correlated with programmed death-ligand (PD-L)1.

Conclusion: To our knowledge, this study is the first to analyze the relationship between an unbiased microbiome-driven cluster and clinical phenotype in children with asthma. The cluster dominated by Haemophilus and Neisseria showed fixed airflow obstruction and mixed granulocytic asthma, which correlated with PD-L1 levels. Thus, microbiome-driven unbiased clustering can help identify new asthma phenotypes related to endotypes in childhood asthma.

RESUMEN

Antecedentes: Los estudios del microbioma respiratorio han favorecido nuestra comprensión de diversos fenotipos y endotipos del asma. Sin embargo, la relación entre el microbioma respiratorio y los fenotipos clínicos en niños con asma sigue sin estar clara. Nuestro objetivo fue identificar, en niños con asma, agrupaciones (clúster) de microbiomas que identifiquen las características clínicas del asma y sus microbiotas dominantes.

Métodos: Se recogió esputo inducido de niños con asma y se generaron perfiles de microbioma mediante secuenciación de la región V3-V4 del gen 16S rRNA. El análisis de clúster se realizó usando el algoritmo PAM (Partitioning Around Medoids). El microbiota dominante en cada clúster se determinó mediante el análisis lineal discriminante. En cada conglomerado se analizó la asociación entre el microbiota dominante, el fenotipo clínico y la citocina inflamatoria.

Resultados: Se evaluaron 83 niños diagnosticados de asma. Entre los cuatro clústeres que reflejaban las características clínicas del asma, el clúster 1, dominado por Haemophilus y Neisseria, se caracterizaba por tener un volumen espiratorio forzado en 1 segundo (FEV1) y la capacidad vital forzada (FVC), posbroncodilatador (BD) inferior al de los demás clúster y un asma granulocítica más mixta. Neisseria se correlacionó negativamente con el VEF1/CVF pre y post-BD. Haemophilus y Neisseria se correlacionaron positivamente con el ligando de muerte programada (PD-L)1.

Conclusiones: Hasta donde sabemos, este estudio es el primero en analizar la relación entre un clúster no sesgado de microbioma y el fenotipo clínico en niños con asma. El clúster dominado por Haemophilus y Neisseria mostró obstrucción fija del flujo aéreo y asma granulocítica mixta, que se correlacionó con los niveles de PD-L1. Así pues, la agrupación no sesgada derivada del análisis del microbioma puede ayudar a identificar nuevos fenotipos de asma relacionados con los endotipos en el asma infantil.

SUMMARY BOX

– What do we know about this topic?
The microbiome-driven unbiased clustering analysis in childhood asthma suggested that a cluster primarily composed of *Haemophilus* and *Neisseria* displayed a fixed airflow obstruction and mixed granulocytic asthma. This observation suggested a possible connection to programmed death-ligand 1.

– How does this study impact our current understanding and/or clinical management of this topic?
The new asthma endotyping driven by airway microbiome provides valuable information for an elaborate classification of clinically heterogenous asthma. It could enable us to determine precise management modalities and predict prognosis in children with asthma.
INTRODUCTION

Identifying various asthma phenotypes and endotypes facilitates a more systemic and differentiated approach for efficient and personalized treatment of asthma, which exhibits heterogeneity and represents the “syndrome” instead of a single simple disease [1]. Phenotype characterizes the outward clinical features, including the inflammatory cell type and airway obstruction or reversibility; it can be applied intuitively in clinics [2]. Conversely, endotype provides a comprehensive understanding of the underlying biological mechanisms at the molecular level, including cytokine or microbiome profiling, which can be used to identify disease-specific markers [3]. Identifying the relationship between phenotype and endotype helps predict the prognosis of heterogeneous asthma and determine the course of treatment [1].

The sputum inflammatory marker, a characteristic of asthma phenotype, can be used as a representative tool to understand and explain the diversity and heterogeneity of asthma [4]. Eosinophilic inflammation induced by heightened T helper 2 (Th2) immune response has been suggested as a classical hypothesis of asthma, whereas neutrophil inflammation is characteristic of non-atopic asthma, which is resistant to steroids [5,6]. The respiratory microbiome is an important tool for determining asthma endotype to help understand the underlying mechanism of asthma development and exacerbation, which may
be related to the sputum inflammatory phenotype [7,8]. Early asymptomatic *Streptococcus* colonization was suggested as a strong predictor of asthma development [9]. The gram-negative microbes or airway microbiome composition and diversity could be related to asthma exacerbation [10]. Respiratory microbiome diversity is reduced in neutrophilic asthma, and opportunistic microbes, such as the genus *Haemophilus*, are replaced, which could be related to severe asthma [11]. An unbiased clustering of the microbiome may reflect the clinical characteristics and severity of asthma [12].

Unlike adult asthma, childhood asthma shows a distinct feature of allergic comorbidities, including atopic dermatitis or food allergy relating to the allergic march, presumably related to the Th2 immune response and eosinophilic inflammation in most cases [13]. However, neutrophilic asthma has recently been reported in majority of the children, which might primarily be due to bacterial and/or viral infection [14]. Owing to the limitation of sampling in children compared with those in adults, limited studies have assessed the relationship between the respiratory microbiome and clinical phenotypes in children with asthma [15].

Therefore, we aimed to classify and characterize the respiratory microbiome in children with asthma using unbiased clustering methods and evaluate the relationship of these microbiome features with clinical phenotypes, including sputum inflammatory phenotype,
bronchial hyperresponsiveness (BHR), bronchodilator responsiveness (BDR), and airway obstruction. We also aimed to evaluate inflammatory cytokines to elucidate the mechanisms by which microbiome-driven inflammation can affect distinct phenotypes.

METHODS

Participants

We screened children who visited the Severance Children’s Hospital for work-up or treatment of asthma from January 2015 to December 2018. The children underwent spirometry, sputum induction, and blood sampling at the first visit, followed by the provocation challenge test at the second visit.

Children with typical asthmatic symptoms, such as recurrent cough or dyspnea, shortness of breath, and chest tightness, underwent spirometry with a bronchodilator (BD) and the bronchoprovocation test. Asthma was diagnosed based on the Global Initiative for Asthma guidelines if a 20% reduction in the forced expiratory volume in 1 second (FEV1) occurred in response to a provocative concentration of inhaled provocholine (PC20 < 10 mg/mL) or BD response, which was verified as a >12% increase in FEV1 after inhaling 200 μg albuterol [16]. We excluded children with the following symptoms: 1) fever, myalgia, purulent sputum, persistent wet cough, and runny and congested nose for 10 days, which
occur in differential diagnoses of asthma, including acute respiratory infection; and 2) cough when feeding or vomiting easily, which occur in cardiac murmur or aspiration [16]. Children with acute asthma exacerbation in the previous four weeks requiring systemic corticosteroid administration or increased use of inhaled corticosteroids were also excluded [17].

Specific serum immunoglobulin E (IgE) levels for the following common inhalant allergens in Korea were measured using the Pharmacia CAP assay (Uppsala, Sweden): two types of dust mites (*Dermatophagoides pteronyssinus* and *Dermatophagoides farina*); cat and dog epithelium; cockroach; mold; and pollen allergens, including *Alternaria*, birch, mugwort, Japanese hop, and ragweed. Atopy was defined as ≥ 0.35 KUa/L of specific IgE for more than one allergen.

**Sputum induction and processing**

After washing their mouths thoroughly with water, all children inhaled 3% saline solution nebulized in an ultrasonic nebulizer (NE-U12; Omron Co., Tokyo, Japan) at maximum output at room temperature and were encouraged to cough deeply at 3-min intervals thereafter. For cell count and microbiome analysis, sputum samples were stored at 4 °C for no more than 2 h before further processing. A fraction of the sample was diluted with phosphate-buffered saline (PBS) containing 10 mmol/L dithiothreitol (WAKO Pure
Chemical Industries Ltd, Osaka, Japan). For cytokine analysis, another fraction of the sample was gently vortexed at room temperature for 20 min after diluting with PBS containing 10 mmol/L dithiothreitol. Sputum aliquots for microbiome and cytokine analysis were stored at -20 °C immediately after collection and then at -70 °C within 12 h to maintain acceptable quality for microbiome analysis [18].

Sputum samples were classified as eosinophilic (>2.5% eosinophils), neutrophilic (>54% neutrophils), mixed granulocytic (>2.5% eosinophils, >54% neutrophils), or paucigranulocytic (≤2.5% eosinophils, ≤54% neutrophils). [19]

This study was approved by the Institutional Review Board of Severance Hospital (protocol no. 4-2004-0036). Written informed consent was obtained from the participants and their parents.

**DNA extraction, PCR amplification, and sequencing**

DNA extraction, PCR amplification, and sequencing were performed concurrently for all samples stored at -70 °C during the recruitment period from 2015–2018. For microbiome analysis, total DNA was extracted from a fraction of the sputum sample using the FastDNA® SPIN Kit for Soil (MP Biomedicals, USA) in accordance with the manufacturer’s instructions. The ratio of absorbance was calculated at 260 nm and 280 nm (A260/A280) to assess the
purity of DNA. The A260/A280 values of all samples were >2.0, indicating that the purity of DNA was acceptable [20]. PCR amplification was performed using fusion primers targeting the V3–V4 regions of the 16S rRNA gene with the extracted DNA. For bacterial amplification, fusion primers of 341F (5′-AATGATACGGCGACCACCGAGATCTACAC-XXXXXXXXXTCGTCGGCAGCGTC-AGATGTGTATAAGAGACAG-CCTACGGGNGGCWGCAG-3′; underlined sequence indicates the target region primer) and 805R (5′-CAAGCAGAAGACGGCATACGAGATGTCTCTCGTGGGCTCGG-AGATGTGTATAAGAGACAG-GACTACHVGGGTATCTAATCC-3′) were constructed in the following order: P5 (P7) graft binding, i5 (i7) index, Nextera consensus, Sequencing adaptor, and Target region sequence.

Amplifications were performed under the following conditions: initial denaturation at 95 °C for 3 min, followed by 25 cycles of denaturation at 95 °C for 30 s, primer annealing at 55 °C for 30 s, extension at 72 °C for 30 s, and final elongation at 72 °C for 5 min.

The PCR product was confirmed using 1% agarose gel electrophoresis and visualized using a Gel Doc system (BioRad, Hercules, CA, USA). The amplified products were purified with Clean PCR (CleanNA). Equal concentrations of purified products were pooled together, and short fragments (non-target products) were removed using Clean PCR (CleanNA). The quality and product size were assessed using a Bioanalyzer 2100 (Agilent, Palo Alto, CA,
USA) with a DNA 7500 chip. Mixed amplicons were pooled and sequenced at Chunlab, Inc. (Seoul, Korea), with Illumina MiSeq Sequencing system (Illumina, USA), according to the manufacturer’s instructions.

**Microbiome data analysis**

Raw reads were processed by performing a quality check and filtering low-quality (<Q25) reads using Trimmomatic ver. 0.32 [21]. After completing the quality check, paired-end sequence data were merged using the fastq_mergepairs command of VSEARCH version 2.13.4 [22] with default parameters. Primers were trimmed using the alignment algorithm of Myers & Miller [23] at a similarity cut-off of 0.8. Non-specific amplicons that did not encode the 16S rRNA were detected using nhmmer [24] in the HMMER software package ver. 3.2.1 with hmm profiles. Unique reads were extracted, and redundant reads were clustered with unique reads using the derep_full_length command of VSEARCH [22]. The EzBioCloud 16S rRNA database [25] was used for the taxonomic assignment using the usearch_global command of VSEARCH [22], followed by more precise pairwise alignment [23]. Chimeric reads were filtered based on <97% similarity by reference-based chimeric detection using the UCHIME algorithm [26] and the non-chimeric 16S rRNA database from EzBioCloud. After chimeric filtering, reads that were not identified to the species level (with <97% similarity) in
the EzBioCloud database were compiled, and de-novo clustering was performed using the cluster_fast command [22] to generate additional operational taxonomic units (OTUs). OTUs with single reads (singletons) were omitted from further analysis.

**Cytokine analysis**

Cytokine analysis of sputum was performed using a human fixed immunotherapy discovery magnetic panel-24 plex kit (Magnetic Luminex® Performance Assay multiplex kit, R&D Systems, Minneapolis, MN, USA). This kit was used to analyze cluster of differentiation (cd)40, granulocyte-macrophage colony-stimulating factor, granzyme B, interferon-α, interferon-γ, interleukin (IL)-1α, IL-1β, IL-1RA, IL-2, IL-4, IL-6, IL-8, IL-10, IL-12p70, IL-13, IL-15, IL-17A, IL-33, C-X-C motif chemokine 10, monocyte chemoattractant protein-1, macrophage inflammatory proteins (MIP)-1α, MIP-1β, programmed death-ligand (PD-L)1, and tumor necrosis factor-α.

**Microbiome data analysis for the clustered groups**

Samples were clustered using species-level abundance data with partitioning around medoid (PAM) clustering based on Jensen–Shannon divergence [27]. The Calinski–Harabasz (CH) index was calculated according to the number of clusters and used to determine the
optimal number of clusters [28]. The resulting clusters were visualized using R with package “ade4” for principal coordinate analysis (PCoA) based on Jensen–Shannon divergence [29].

Linear Discriminant Effect Size (LEfSe) analysis was performed to discover microbiota as a biomarker related to each cluster. The clusters showed a significant difference in the analysis using the Kruskal–Wallis test. Significant biomarkers were obtained with linear discriminant analysis (LDA) score >4.0 and $p$-value < 0.05 in the pairwise comparison using Mann–Whitney test and Bonferroni’s methods [30]. The resulting biomarkers were visualized using GraPhlAn for cladogram and R statistical package (R version 3.2.5.; Institute for Statistics and Mathematics, Vienna, Austria; www.R-project.org) with package “ggplot” for boxplot using the Kruskal–Wallis H test [31].

**Statistical analyses**

The clusters were defined using PAM clustering and the CH index. For evaluating clinical characteristics across the clusters, we compared the participants’ demographics, the pulmonary function parameters, such as airway obstruction index (FEV1, FEV1/forced vital capacity (FVC)), fixed airway obstruction index (post-BD FEV1, post-BD FEV1/FVC), AHR and BD response, and sputum inflammatory phenotype across the clusters. Student’s t-test, Mann–Whitney test, one-way analysis of variance, or Kruskal–Wallis test was used for
continuous variables. Chi-Squared or Fisher’s exact test was used for categorical variables.

Post-hoc analysis with Bonferroni correction was performed if a significant difference was observed between the four clusters. Spearman’s rank correlation was used to assess the relationship between the microbiota as a biomarker for the clusters vs. inflammatory cytokines and pulmonary function parameters. *P*-values <0.05 were considered statistically significant. SPSS version 23 statistical software (SPSS, Inc., Chicago, IL, USA) and R statistical package (R version 3.2.5.; Institute for Statistics and Mathematics, Vienna, Austria; www.R-project.org) were used for analysis.

**RESULTS**

**Clinical characteristics across the clusters**

Eighty-three children diagnosed with asthma (median age: 7.5 years, 31.3% boys) were evaluated. The majority of the children, approximately 83%, were atopic.

According to the number of clusters, defined using the PAM clustering method, a higher CH index was obtained in two and four clusters than in other numbers of clusters (Figure 1A), which were well-separated in the PCoA plots (Figures 1B–D).

We compared the clinical characteristics of the participants across the two and the four clusters. The four clusters showed some significantly different clinical characteristics,
including inflammatory phenotype (p = 0.007) and pulmonary function parameters (Table 1), whereas the two clusters did not show any significantly different clinical characteristics (Supplementary Table 1). Post-BD FEV1/FVCs (p = 0.020) differed significantly across the four clusters; however, the difference in pre-BD FEV1/FVC (p = 0.060) was not statistically significant. Therefore, we comprehensively evaluated the clinical characteristics and microbiome profile in the four clusters to identify microbiotas as meaningful biomarkers related to clinical characteristics, such as inflammatory phenotype and pulmonary function parameters.

**Sputum inflammatory phenotype and pulmonary function parameters across the clusters**

Post-hoc analysis was performed to identify the significantly different clusters in inflammatory phenotype. Only clusters 1 and 2 exhibited a significant difference with multiple corrections (Figure 2A). Post-hoc analysis was also performed to identify which inflammatory phenotype differed significantly in clusters 1 and 2. Since there was no significantly different inflammatory phenotype, which could explain the difference between clusters 1 and 2, the difference in inflammatory phenotype among these two clusters was evaluated without Bonferroni correction (Figure 2B). This explorative investigation revealed
differences in the mixed granulocytic and paucigranulocytic types in clusters 1 and 2 (Figure 2B).

Cluster 1 had a lower post-BD FEV1/FVC than those in the other clusters (Figure 2C). In a pairwise comparison between two clusters, post-BD FEV1/FVC of cluster 1 was significantly lower than that of cluster 2 \((p = 0.031)\) after Bonferroni’s correction.

In summary, cluster 1 had a lower post-BD FEV1/FVC, indicating fixed airflow obstruction and more mixed granulocytic and paucigranulocytic asthma.

**Dominant microbiotas in the clusters**

Since the 16S rRNA analytic method has limitations in identifying an individual microbe at the species level when applied with only partial amplicons [8], the abundance of the microbiotas was analyzed up to the genus level (Supplementary Fig. 1), and compared at the genus level among the clusters (Figure 3) at \(p < 0.05\) using the Kruskal–Wallis H test to identify the dominant microbiotas related to each cluster. Microbiotas were selected at the genus level with an LDA score >4.0, as seen in the LDA histogram and cladogram in Figure 4. The predominance was as follows: *Neisseria* and *Haemophilus* in cluster 1; *Prevotella*, *Veillonella*, and *Actinomyces* in cluster 2; *Streptococcus* and *Granulicatella* in cluster 3; and *Ralstonia* in cluster 4.
Correlation between microbiota vs. inflammatory cytokines and pulmonary function

The correlation between the prominent genera and inflammatory cytokines was analyzed (Supplementary Table 2). Among the 83 participants, samples from 63 participants were available for analyzing inflammatory cytokines. Since cluster 1 had a more mixed granulocytic type and fixed airway obstruction, and *Neisseria* and *Haemophilus* were predominant in cluster 1, we focused on the cytokine that showed a significant correlation with these two genera. Only PD-L1 had a meaningful correlation with both microbes ($r = 0.445, p = 0.016$ for *Neisseria*; $r = 0.450, p = 0.014$ for *Haemophilus*).

The correlation between the predominant genus, including *Neisseria* and *Haemophilus*, and the less abundant genus, including *Streptococcus*, in cluster 1 vs. pre-BD and post-BD FEV1/FVC indices were analyzed (Figure 5). Only *Neisseria* correlated negatively with pre-BD FEV1/FVC ($r = -0.227, p = 0.039$) and post-BD FEV1/FVC ($r = -0.227, p = 0.039$), whereas the other microbiotas showed no significant correlation with the pre-BD and post-BD FEV1/FVC indices.

DISCUSSION

An unbiased microbiome profile clustering method used in children with asthma revealed that the cluster with abundant *Neisseria* and *Haemophilus* exhibited fixed airflow obstruction based on the post-BD FEV1/FVC index and more mixed granulocytic phenotype.
The pre-BD and post-BD FEV1/FVC indices decreased with an increase in the relative abundance of *Neisseria*, indicating that *Neisseria* could be related to airway obstruction in childhood asthma. *Neisseria* and *Haemophilus* correlated positively with PD-L1 levels, suggesting that they could affect fixed airflow obstruction and mixed granulocytic phenotype in relation to PD-L1 in childhood asthma.

Microbiome study is used for asthma endotyping, which defines the subtypes of heterogenous asthma based on the underlying pathologic mechanisms [7]. Previous studies on microbiome data are limited to a supervised approach using known clinical phenotypes and could not address independent microbiome-driven subtyping [11,32,33]. A recent microbiome study in adult asthma suggested the clinical significance of unbiased clustering based on microbiome profiles alone [12]. We applied this unbiased clustering method in children with asthma, and the cluster showed a significant association with clinical characteristics, including fixed airflow obstruction and mixed granulocytic type. Thus, the unbiased cluster analysis of airway microbiome was clinically meaningful in childhood asthma.

*Haemophilus*, a pathogenic microbe found in airway dysbiosis, is considered a major pathogenic microbiota in asthma attacks [10]. It is highly abundant in the neutrophilic phenotype of severe asthma [11] and is prominent in eosinophilic asthma [33]. The relevance of *Neisseria* in eosinophilic asthma is debatable [32,33]. As both neutrophilic and...
eosinophilic inflammatory processes play a role in asthma related to Th1 and Th2 immune response, and many debatable results have been reported [13,14], it is reasonably acceptable that cluster 1 has more mixed granulocytic asthma.

PD-L1, showing a significant positive correlation with *Neisseria* and *Haemophilus* in our study, may strengthen Th2 inflammation and increase AHR in asthma; however, it can suppress CD8 T-cell immunity, preventing the clearance of infected pathogens from the perspective of acute infection [34,35]. The dual roles of PD-L1 in asthma, including strengthening Th2 inflammation and weakening innate immunity from infected pathogens, can explain its contribution to asthma exacerbation [10]. These dual roles can also contribute to the eosinophilic inflammation through Th2 immune response and the neutrophilic inflammation through recurrent infection. It can cause a more mixed phenotype in the cluster in which *Neisseria* and *Haemophilus* were dominant in our study.

There are few studies on fixed airflow obstruction in children, a characteristic of chronic obstructive pulmonary disease (COPD), which can be an index of severe asthma when accompanied by asthma in adults [36-38]. It generally develops owing to airway remodeling driven by chronic inflammation [36,39]. Frequent asthma exacerbation can be a risk factor for fixed airflow obstruction in children with asthma [40], and infections are the leading cause of asthma exacerbation in children [41,42]. In this study, cluster 1 showed a
mixed granulocytic phenotype, causing fixed airflow obstruction owing to increased inflammation reactions triggered by eosinophilic and neutrophilic inflammation responses [43]. This finding is supported by previous reports revealing that overlapping inflammatory pathways, presenting as elevated eosinophils and neutrophils, might be detrimental to lung function loss [44].

*Neisseria* and *Haemophilus* were predominant in cluster 1, which showed fixed airflow obstruction; in contrast, *Prevotella*, *Veillonella*, and *Actinomyces* were predominant in cluster 2, which showed favorable lung function. This finding is in line with that of previous reports showing that airway microbial dysbiosis with the overgrowth of opportunistic pathogens and lesser normal airway microbes can simultaneously develop and aggravate asthma [11]. *Neisseria* correlated independently with airflow limitation parameters, similar to previous findings; the increased prevalence of *Neisseria* owing to rhinovirus infection can induce the immunomodulatory properties of dendritic cells and proinflammatory cytokines [45,46], which might affect pulmonary function. This possible explanation is justified in children, who are increasingly exposed to respiratory viral infections [47].

Cluster 2, showing favorable lung function, predominantly included *Prevotella*, *Veillonella*, and *Actinomyces* in our study. *Prevotella* is more predominant in controls and infants without wheezing than in patients and infants with asthma or COPD and wheezing
Prevotella was suggested to reduce pathologic Haemophilus influenza-induced IL-12p70 [49] and neutrophilic airway inflammation [50]. However, the presence of Prevotella and Veillonella at one month of age was associated with the incidence of asthma at six years of age [51]. Actinomyces were less abundant in acute asthma exacerbation than in stable asthma [10]. It is also less abundant in neutrophilic asthma, which is considered a severe type of asthma [11].

Cluster 3 included older patients and more women than the other clusters. Streptococcus and Granulocatella were predominant in cluster 3. As reported previously, Streptococcus was most abundant in our study [52], and it is an early marker for predicting asthma development during later childhood in infants [9]. In contrast, cluster 4 included younger patients, and Ralstonia was predominant in cluster 3. Ralstonia, considered a pathologic Pseudomonas until recently, was reported to be positively correlated with pyruvic acid, which has a crucial protective role in IgE production in response to allergens [53].

Airway microbiomes exhibit distinct features according to age and sex; however, this finding has not been adequately addressed [52,54,55].

This study has several limitations. First, the number of patients for evaluation of the four clusters was small. Second, we could not evaluate the two clusters with the most optimal CH index as these clusters could not explain the clinical characteristics. Third, we could not
collect detailed clinical information, including the degree of control in asthma, asthma duration, drug usage, and the frequency of acute exacerbation. Despite these limitations, to the best of our knowledge, this study is the first to analyze the relationship between an unbiased microbiome-driven cluster and clinical phenotype in children with asthma. In addition, it is meaningful that the characteristics of fixed airflow obstruction and mixed granulocytic asthma in children, which were sporadically reported, were assessed through the relationship between the microbiome and inflammatory cytokines. The findings of this study provide insights into the effect of the airway microbiome on lung function, which has not been addressed [42].

In conclusion, the microbiome-driven unbiased clustering method in childhood asthma can help find new endotype-related asthma phenotypes. Our findings suggested that the cluster dominated by *Haemophilus* and *Neisseria* found through this method shows fixed airflow obstruction and mixed granulocytic asthma, which can be related to PD-L1. Thus, new asthma endotyping driven by airway microbiome can provide valuable information for determining precise management modalities and predicting prognosis in children with asthma.
Funding

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant number: NRF-2021R1C1C1011294).

Conflicts of interest

The authors have no conflicts of interest to declare.
REFERENCES


Inflammatory phenotypes in patients with severe asthma are associated with distinct 

Sputum microbiome profiles identify severe asthma phenotypes of relative stability at 

13. Douwes J, Gibson P, Pekkanen J, Pearce N. Non-eosinophilic asthma: importance and 

14. McDougall CM, Helms PJ. Neutrophil airway inflammation in childhood asthma. 

15. Shah R, Bunyavanich S. The airway microbiome and pediatric asthma. Curr Opin 

Global strategy for asthma management and prevention: GINA executive summary. 
Eur Respir J. 2008;31:143-78.

17. Fergusson JE, Patel SS, Lockey RF. Acute asthma, prognosis, and treatment. J Allergy 

between collection and storage significantly influences bacterial sequence 
composition in sputum samples from cystic fibrosis respiratory infections. J Clin 


20. Koetsier G, Cantor E. A practical guide to analyzing nucleic acid concentration and 


### Table 1. Patient characteristics across the four clustered groups (N = 83)

<table>
<thead>
<tr>
<th></th>
<th>Total (n = 83)</th>
<th>Cluster 1 (n = 15)</th>
<th>Cluster 2 (n = 39)</th>
<th>Cluster 3 (n = 16)</th>
<th>Cluster 4 (n = 13)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age, years</strong></td>
<td>7.5 (6.5–9.7)</td>
<td>8.2 (6.5–10.4)</td>
<td>7.5 (6.3–9.7)</td>
<td>8.9 (7.5–10.7)</td>
<td>6.5 (5.7–7.9)</td>
<td>0.020</td>
</tr>
<tr>
<td><strong>Male sex, n (%)</strong></td>
<td>26 (31.3)</td>
<td>14 (93.3)</td>
<td>24 (61.5)</td>
<td>9 (56.3)</td>
<td>10 (76.9)</td>
<td>0.081</td>
</tr>
<tr>
<td><strong>Atopy, n (%)</strong></td>
<td>29 (82.9)</td>
<td>10 (66.7)</td>
<td>31 (79.5)</td>
<td>15 (93.8)</td>
<td>11 (84.6)</td>
<td>0.282</td>
</tr>
<tr>
<td><strong>Sputum inflammatory phenotype, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eosinophilic</td>
<td>27 (32.5)</td>
<td>3 (20.0)</td>
<td>18 (46.2)</td>
<td>1 (6.3)</td>
<td>5 (38.5)</td>
<td></td>
</tr>
<tr>
<td>Neutrophilic</td>
<td>32 (38.6)</td>
<td>2 (13.3)</td>
<td>15 (38.5)</td>
<td>10 (62.5)</td>
<td>5 (38.5)</td>
<td>0.007</td>
</tr>
<tr>
<td>Mixed</td>
<td>13 (15.7)</td>
<td>6 (40.0)</td>
<td>4 (10.3)</td>
<td>2 (12.5)</td>
<td>1 (7.7)</td>
<td></td>
</tr>
<tr>
<td>Paucigranulocytic</td>
<td>11 (13.3)</td>
<td>4 (26.7)</td>
<td>2 (5.1)</td>
<td>3 (18.8)</td>
<td>2 (15.4)</td>
<td></td>
</tr>
<tr>
<td><strong>Pulmonary function parameters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEV1, % predicted</td>
<td>96.8 ± 16.2</td>
<td>92.4 ± 16.0</td>
<td>98.4 ± 16.2</td>
<td>98.5 ± 15.5</td>
<td>94.8 ± 17.8</td>
<td>0.610</td>
</tr>
<tr>
<td>FEV1/FVC</td>
<td>81.4 (74.1–85.5)</td>
<td>0.76 (0.70–0.81)</td>
<td>0.83 (0.74–0.86)</td>
<td>0.84 (0.80–0.88)</td>
<td>0.80 (0.71–0.84)</td>
<td>0.060</td>
</tr>
<tr>
<td>Post BD FEV1, % predicted</td>
<td>105.7 ± 15.8</td>
<td>101.3 ± 18.5</td>
<td>108.1 ± 15.1</td>
<td>105.1 ± 14.6</td>
<td>104.3 ± 16.2</td>
<td>0.540</td>
</tr>
<tr>
<td>Post BD FEV1/FVC</td>
<td>85.3 (80.8–91.1)</td>
<td>80.5 (75.8–86.3)</td>
<td>88.0 (83.0–91.3)</td>
<td>86.5 (83.0–91.8)</td>
<td>84.5 (77.3–89.0)</td>
<td>0.020</td>
</tr>
<tr>
<td>BDR assessing Δ FEV1</td>
<td>29 (34.9)</td>
<td>7 (46.7)</td>
<td>14 (35.9)</td>
<td>4 (25.0)</td>
<td>4 (30.8)</td>
<td>0.633</td>
</tr>
<tr>
<td>BHR assessing challenge test</td>
<td>62 (74.7)</td>
<td>8 (57.1)</td>
<td>32 (84.2)</td>
<td>12 (75.0)</td>
<td>10 (83.3)</td>
<td>0.204</td>
</tr>
</tbody>
</table>

FEV1, forced expiratory volume in one second; FVC, forced vital capacity; BD, bronchodilator; BDR, bronchodilator response; Δ, the change of before and after bronchodilator; BHR, bronchial hyperresponsiveness response.
FIGURE LEGENDS

Fig. 1. (A) Calinski–Harabasz (CH) index according to cluster number using partitioning around medoid clustering method based on Jensen–Shannon divergence at the species level. (B) Two-dimensional (2D) principal coordinate analysis (PCoA) plot for cluster 2. (C) 2D PCoA plot for cluster 4. (D) Three-dimensional PCoA plot for cluster 4.
Fig. 2 (A) Comparison of the sputum inflammatory phenotype across the clusters. (B) Comparison of the sputum inflammatory phenotype in clusters 1 and 2. (C) Post-bronchodilator (BD) forced expiratory volume in 1 second (FEV1)/forced vital capacity (FVC) across the clusters. The P value was calculated using post-hoc analysis with Bonferroni correction. *P was calculated using post-hoc analysis without Bonferroni correction for exploration.
**Fig. 3** Comparison of the microbiotas at the genus level among the clusters with $p<0.001$ in the Kruskal–Wallis H test. The red lines represent $p<0.05$ in the pairwise comparison using Mann–Whitney test and Bonferroni correction.
Fig. 4 (A) Linear discriminant analysis (LDA) effect size analysis across the four clusters with \( p < 0.05 \) and an LDA score >4.0 (B) Cladogram showing differentially abundant taxa according to each cluster. *Haemophilus parainfluenzae* group (*Haemophilus influenza*, *Haemophilus aegyptius*, and unclassified microbes), *Streptococcus salivarius* group (*Streptococcus salivarius* subsp. *salivarius*, *Streptococcus thermophiles*, *Streptococcus vestibularis*, and unclassified microbes), *Streptococcus sinensis* group (*Streptococcus sinensis* and unclassified microbes), *Streptococcus pneumoniae* group (*Streptococcus pneumonia*, *Streptococcus oralis* subsp. *oralis*, *Streptococcus oralis* subsp. *tigurinus*, *Streptococcus oralis* subsp. *dentisani*, *Streptococcus mitis*, *Streptococcus infantis*, *Streptococcus pseudopneumoniae*, *Streptococcus timonensis*, and unclassified microbes), *Streptococcus sanguinis* group (*Streptococcus sanguinis* and unclassified microbes), and *Granulicatella adiacens* group (*Granulicatella adiacens* and unclassified microbes).
**Fig. 5** Correlation between pre-BD and post-BD FEV1/FVC vs. *Neisseria*

BD, bronchodilator; FEV1, forced expiratory volume in 1 second; FVC, forced vital capacity.

**Supplementary Fig. 1** Abundance of the microbiota in the four clusters at each taxonomic categorical level; (A) phylum, (B) order, (C) class, (D) family, and (E) genus.