Extracellular vesicle-derived microbiome obtained from exhaled breath condensate in patients with asthma

Jin An[1], Andrea McDowell[2], Yoon-Keun Kim[2], and Tae-Bum Kim[3]

[1]Kyung Hee University Hospital at Gangdong
[2]MD healthcare, Inc
[3]Asan Medical Center

March 30, 2022

Abstract

BACKGROUND: Despite the construction of the metagenome of the asthmatic lung, limitations persist in sampling the bronchial airway. This study analyzed extracellular vesicles (EVs) obtained from exhaled breath condensate (EBC) to compare the distinct characteristics of the microbiome in asthmatics with those in healthy controls and proposed a diagnostic artificial intelligence-based model of asthma. METHODS: We obtained the EBC from 58 healthy controls and 251 patients with asthma. EVs were isolated from the EBC and analyzed. The extracted 16s rDNA was subjected to next generation sequencing. Taxonomic profiling was conducted for all samples at the genus level. A combination of artificial neural network (ANN) and gradient boosting (GBM) was applied to selective EBC biomarkers. RESULTS: The asthma group exhibited significantly higher alpha diversity based on the results of the Chao1, Shannon, and Simpson indices. The bacterial composition of patients with asthma different from that of the controls. At the genus level, Sphingomonas, Akkermansia, Methylophaga, Acidocella, and Marinobacter were significantly more abundant in patients with asthma. The diagnostic model using GBM and ANN demonstrated good performance with respective areas under the curve of 0.832 and 0.769. Firmicutes and Proteobacteria at the phylum level were common important features between the GBM and ANN asthma models. CONCLUSION: We demonstrated a distinct pattern in the microbiome of patients with asthma, indicating the potential role of microbiome-based diagnosis of asthma. To the best of our knowledge, this was the first study to identify the microbiome in asthma using EBC-derived EVs.

Extracellular vesicle-derived microbiome obtained from exhaled breath condensate in patients with asthma

Jin An[1, 2*], Andrea McDowell[3], Yoon-Keun Kim[3], and Tae-Bum Kim[1]

Short title : Metagenomics using exhaled breath condensate in asthma

[1]Department of Allergy and Clinical Immunology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea

[2]Department of Pulmonary, Allergy and Critical Care Medicine, Kyung Hee University Hospital at Gangdong, College of Medicine, Kyung Hee University, Seoul, Republic of Korea

[3]Institute of MD Healthcare Inc., Seoul, Republic of Korea

Correspondence to: Tae-Bum Kim, MD, PhD

Department of Allergy and Clinical Immunology, Asan Medical Center, University of Ulsan College of Medicine, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Republic of Korea. Tel: 82-2-3010-3280; E-mail: tbkim@amc.seoul.kr

Acknowledgments

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean government (MSIT) (2019M3E5D3073365).

Conflict of interest : None

Word count Abstract: 245/ Text: 3,150

Abstract

BACKGROUND: Despite the construction of the metagenome of the asthmatic lung, limitations persist in sampling the bronchial airway. This study analyzed extracellular vesicles (EVs) obtained from exhaled breath condensate (EBC) to compare the distinct characteristics of the microbiome in asthmatics with those in healthy controls and proposed a diagnostic artificial intelligence-based model of asthma.

METHODS: We obtained the EBC from 58 healthy controls and 251 patients with asthma. EVs were isolated from the EBC and analyzed. The extracted 16s rDNA was subjected to next generation sequencing. Taxonomic profiling was conducted for all samples at the genus level. A combination of artificial neural network (ANN) and gradient boosting (GBM) was applied to selective EBC biomarkers.

RESULTS: The asthma group exhibited significantly higher alpha diversity based on the results of the Chao1, Shannon, and Simpson indices. The bacterial composition of patients with asthma different from that of the controls. At the genus level, *Sphingomonas* , *Akkermansia* ,*Methylophaga* , *Acidocella* , and *Marinobacter* were significantly more abundant in patients with asthma. The diagnostic model using GBM and ANN demonstrated good performance with respective areas under the curve of 0.832 and 0.769. Firmicutes and Proteobacteria at the phylum level were common important features between the GBM and ANN asthma models.

CONCLUSION: We demonstrated a distinct pattern in the microbiome of patients with asthma, indicating the potential role of microbiome-based diagnosis of asthma. To the best of our knowledge, this was the first study to identify the microbiome in asthma using EBC-derived EVs.

Keyword: asthma; microbiome; exhaled breath condensate; extracellular vesicles; metagenomics

Abbreviations used

EV - Extracellular vesicles

EBC - Exhaled breath condensate

AI - Artificial intelligence

OUT - Operational taxonomic unit

GBM - Gradients Boosting Machine

ANN - Artificial Neural Network

AUC – Area under the curve

COPD - Chronic obstructive pulmonary diseases

BAL - Bronchoalveolar lavage

Introduction

Asthma is a complex inflammatory disease resulting from the interaction between genetic susceptibility and environmental exposures over the course of the individuals life [1]. The bronchial microbiome of patients with asthma differs from that of the healthy controls [2,3]. Researchers have studied the relationships between the respiratory microbiota and heterogeneity of asthma, and suggested that the microbiome may be a factor

determining asthma progression [4,5]. The potential role of the airway microbiota has been the focus of research on the pathogenesis of asthma.

Metagenomic analysis has been conducted using 16S rDNA sequencing from extracellular vesicles (EVs) to identify microbiomes and provide culture-independent estimation of microbial diversity[6]. EVs, which are nanoparticles secreted by bacteria, have an important biological function as intercellular signaling mediators via enclosed proteins, nucleic acids, and lipids, and as biomarkers for the diagnosis and prognosis of diseases[7,8]. They also contain fragments of bacterial genomic DNA and induce the host immune response, and may reflect the microbiota of the host[9]. Thus, EVs are valuable diagnostic tool for metagenomics.

There is no consensus regarding to the best sample procedure for the metagenomics analysis of respiratory disease. Studies of the lower airways can be challenging because of the limitations associated with sampling the bronchial airway, which involves invasive bronchoscopy. Exhaled breath condensate (EBC) contains aerosol and volatile compounds that can be analyzed to noninvasively understand the physiologic and pathologic processes in the lung[10].

This study performed metagenomic analysis using EBC-dereived EVs to determine and compare the characteristics of the microbiome in patients with asthma with those in healthy controls, and to investigate the microbiome differences between eosinophilic and non-eosinophilic asthma. Moreover, we suggested a diagnostic tool for asthma using artificial intelligence (AI) modeling based on the results of the microbiota composition in patients with asthma.

Methods

Study population

The EBC was obtained from 58 healthy controls and 251 patients with asthma at the Asan Medical Center between September 2014 and December 2019. Healthy controls were screened for respiratory diseases and recruited during regular wellness examinations conducted at Asan Medical Center if they did not have any current or previous history of pulmonary disease, as determined by the medical examination. The eligibility criteria for patients with asthma were as follows[11,12]: (1) patients aged >18 years; (2) symptoms such as dyspnea, wheezing, or cough for more than 3 months; and (3) airway hyperresponsiveness as indicated by a 20% reduction in forced expiratory volume in 1 s (FEV1) with methacholine 16 mg/mL (PC20) on a provocation test, or airway reversibility with FEV1 >12% (and at least 200 mL) after inhalation of a short-acting beta-agonist. Patients with severe lung damage, bronchiectasis, or a history of lung resection were excluded. The participants' EBC samples were collected using the RTube^{TM} (Respiratory Research, Austin, USA). Informed consent was obtained from all participants, and the study protocol was approved by the Institutional Review Board of Asan Medical Center (IRB No. 2006-0388).

Sample processing and 16S rDNA sequencing

The collected EBC samples were stored in R-tubes at -20°C until sample processing for metagenomic sequencing. Due to the relatively low sample volume, 1x phosphate buffered saline solution (Welgene, Gyeongsan, Korea) was added to all EBC samples to attain a total volume of 1 mL for each sample. Subsequently, samples were centrifuged for 10 min (10,000 × g, 4 °C), the supernatant was collected, and passed through a 0.22-μm filter to sterilize the solution by removing any cells or unnecessary particles. The filtrate solution was subjected to a temperature of 100°C for 40 min and centrifuged (13,000 rpm, 4°C) for 30 min to extract DNA from the double membrane-enclosed EVs. The DNA within the subsequent supernatant was extracted according to the instructions provided in the DNeasy PowerSoil kit (QIAGEN, Germany) and quantified using the QIAxpert (QIAGEN, Germany) software. The V3-V4 hypervariable regions of the 16S rDNA contained in each sample were amplified and the libraries were prepared and quantified as described previously[13]. All resulting 16S rDNA amplicons were sequenced using the MiSeq (Illumina, USA).

Taxonomic profiling of EBC samples

The 16S-based microbiome taxonomic profile of the microbial EVs isolated from the participants' EBC

samples was analyzed using MDx-Pro ver. 2 (MD Healthcare, Korea) as described previously [14]. In summary, after trimming and merging the paired-end reads, only sequence reads with lengths greater than 350 bp and below 550 bp were used for further analysis. Operational taxonomic units (OTUs) were assigned using VSEARCH de novo clustering at the genus level with a similarity threshold of 97%. The Silva 132 sequence database was used for taxonomic assignment of the OTUs in all samples.

Asthma AI modeling and feature importance analysis

Microbiome next generation sequencing (NGS) data were converted to relative composition at the genus level to standardize the input data. A total of 2,035 features were detected from all samples at the genus level. Two general linearized models (GLM) were developed using the t- test and linear discriminant analysis effect size (LEfSe) methods for feature selection. Two algorithms were incorporated for AI modeling, including a gradients-boosting machine (GBM). To apply the GBM algorithm to the tabulated microbiome composition data, machine learning modules in the Python scikit-learn package were applied. The second algorithm was an artificial neural network (ANN), which utilizes regularized 5-layered neural networks and was conducted with the Tensorflow ecosystem using Python[15]. The ensemble algorithm was based on the average values obtained from each prediction model. The data were split into testing and training sets 10 times before training to validate the performance of the models. After splitting the samples, the models were trained based on the training set, followed by validation in the test set. Feature importance was assessed using the testing model to determine the relatively important taxa. Permutation importance was applied using the Scikit-learn package to trace the feature importance in the AI model [16].

Statistical Analysis

Alpha diversity was ascertained using the phyloseq R package to calculate the observed species richness and Chao1 for estimated species richness, and Shannon and Simpson indices for species evenness and richness with the rarefaction set at 1091. Beta diversity testing was performed using principal coordinates analysis (PCoA) for multidimensional scaling of the EBC samples to determine dissimilarity between the healthy control and asthma groups using the basic 'stats' R package. T-tests were performed to assess the statistical difference between the control and asthma groups, where $p < 0.05$ was considered statistically significant. R version 3.6.3 was used for all analyses.

Results

Clinical characteristics of the participants

In the present study, EBC samples were obtained from a total of 309 participants, including 58 healthy controls and 251 patients with asthma (**Table 1**). The average age of the participants in the healthy control and asthma groups were 50.8 (SD 14.4) and 32.4 (SD 7.2) years, respectively. The sex ratio (women/men) of in the healthy control group was 1.6, while that in the asthma group was 2.4, i.e., it was skewed more toward women. The clinical characteristics of the 251 patients with asthma are detailed in **Table S1** .

Alpha and beta diversity

Alpha diversity was assessed based on species richness in addition to species evenness and richness, to determine the baseline differential characteristics of the EV microbiome in the control and asthma EBC samples (**Figure 1A**). The asthma group showed significantly higher alpha diversity ($p < 0.05$) for all four measurements. The observed OTUs and Chao1 estimates showed greater species richness in the asthma EBC samples: Chao1 showed a more moderate increase compared to the observed OTU values. Similarly, the Shannon and Simpson diversity indices both demonstrated higher alpha diversity measures in the asthma group than those in the healthy control group, with the most dramatic increase seen in the Shannon diversity index.

PCoA (based on Bray-Curtis distance of all control and asthmatic patient samples) yielded a primary principal coordinate (PCo1) of 20.13%, while the secondary principal coordinate (PCo2) yielded a value of 8.1% (**Figure 1B**). Generally, EBC samples belonging to the healthy control group were clustered below 0.0

along the PCo1 axis and were evenly distributed along the PCo2 axis. Although there were several asthma patient samples in the same cluster of the control group, several asthmatic samples were tightly clustered around the 0.3 value on the PCo1 axis and 0.0 on the PCo2 axis, with only a few control samples included in the cluster. Diversity assessment for the eosinophilic and non-eosinophilic asthmatic patients' EBC samples yielded no significant differences in any of the four alpha diversity metrics or the PCoA plot for beta diversity (p >0.05) (**Figure S1**).

Microbial EV composition at the genus level

Taxonomic profiling of the microbial EVs isolated from EBC samples was conducted as described above, which revealed that *Acinetobacter* ,*Pseudomonas* , *Staphylococcus* , *Bacteroides* , and*Sphingomonas* were the most prominent genera in the EBC samples (**Figure 2A**). *Acinetobacter* was the most common microbial EV genus in the control (13.0%) and asthmatic (5.4%) EBC samples, followed by *Staphylococcus* (5.2%) and *Pseudomonas* (5.0%) in the healthy control group and *Sphingomonas* (5.0%) and*Bacteroides* (4.0%) in the asthma group. Comparison of the average relative abundance of the 25 most common microbial EV genera in the healthy control and asthma groups using a heat map demonstrated the differential abundance within each sample and between the groups (**Figure 2B**). Furthermore, t-test analysis of the average relative abundance of microbial EV genera yielded multiple taxa that significantly differed between the control and asthma groups (**Figure 2C**). The five most common genera that were enriched in the control group in descending order of proportion were*Acinetobacter* , *Staphylococcus* , *Bifidiobacterium* ,*Blautia* , and *Collinsella* , while *Sphingomonas* ,*Akkermansia* , *Methylophaga* , *Acidocella* , and*Marinobacter* were the most common five genera that were significantly depleted in the healthy control group (p <0.05). The compositional variation between the microbial EV profiles of EBC samples obtained from patients with non-eosinophilic and eosinophilic asthma was also assessed, and no discernable difference was observed at the phylum and genus levels (**Figures S2 and S3**).

EBC-based asthma diagnostic model performance assessment

The 16S rDNA EV metagenomic profiles of the EBC samples were used as features for the various methods for the development of diagnostic models for asthma. Ten iterations of each modeling method confirmed that logistic methods using either t-test of LEfSe biomarker selection performed more poorly than any of the ML methods based on the area under the curve (AUC) values (**Figure 3A**). The incorporation of LEfSe biomarkers as features for logistic models boosted the median AUC value of the t-test method from 0.749 to 0.760; however, t- test feature selection produced the higher average AUC value of the two methods (**Table 2**). While the ANN method demonstrated a higher average AUC value than that of either of the logistic models, GBM's average AUC value of 0.832 was the highest among the five methods, including the combined GBM/ANN ensemble methodology, which yielded a slightly lower average AUC value of 0.826. The standard deviation between the 10 iterations of each method was relatively low, ranging from 0.029 to 0.050. Receiver operating characteristic (ROC) curve plots also depicted the AUC values of the 10 model iterations of each asthma model method based on the range of specificity and sensitivity values (**Figure 3B**).

Common important features between the GBM and ANN asthma models

Unlike the logistic models, each ML method used in this study, i.e., GBM and ANN, incorporated all microbial EV genera as features. Therefore, feature importance evaluation was conducted after developing the ML models to determine the genera with the greatest impact on both asthma diagnostic models. Twenty-one of the 50 highest ranked features in both the GBM and ANN asthma models were shared between both methods (**Table 3**). At the phylum level, Firmicutes accounted for the most significant features, with a total of 9 genera belonging to the Firmicutes phylum, followed by Proteobacteria with 7 significant features. At the genus level, *Ruminococcaceae UCG-014* ,*Lachnospiraceae UCG-008* , *Pseudomonas* ,*Acinetobacter* , *Eubacterium hallii group* , *Blautia* ,*Bifidobacterium* , *Collinsella* , *Paracoccus* , and*Holdemanella* were the 10 most important features in the GBM and ANN asthma models.

Discussion

This was the first study to identify bacterial compositional differences between patients with asthma and healthy controls using EBC-derived EVs. Patients with asthma exhibited significantly different diversity and richness indicators compared to the healthy controls. Five bacterial genera, including *Sphingomonas* , *Akkermansia* ,*Methylophaga, Acidocella* , and *Marinobacter* were significantly abundant in the EBC EVs of patients with asthma, whereas*Acinetobacter* , *Staphylococcus* , *Bifidobacterium* ,*Blautia* , and *Collinsella* were more abundant in the controls. No significant difference was observed between the EBC microbiota in eosinophilic and non-eosinophlic asthma. We suggested a diagnostic tool for asthma using AI modeling based on the distinctness of the bacterial community. Firmicutes and Proteobacteria were common important features at the phylum level in the GBM and ANN asthma models.

This metagenomic analysis of EBC-derevied EVs enhances the understanding of the bacterial communities residing in the airway. EVs are composed of a lipid bilayer that contains transmembrane proteins which affect host cells and other bacteria, even those located far from the colonized area[17,18]. Moreover, they participate in intercellular communication and have potential applications as biomarkers and therapeutic agents for several allergic diseases[19,20]. Airway studies focusing on the microbiota and EVs in nasal lavage and lung tissue samples reported that EVs have distinct characteristics associated with chronic rhinitis and chronic obstructive pulmonary diseases (COPD) [18,21].

Although the advances in metagenomics have elucidated the lung microbiota, sampling the lower respiratory tract is a challenge. Bronchoscopy enables investigation of the lower airway with greater precision, but is invasive. The use of bronchial brushes and bronchoalveolar lavage (BAL) fluid is disputed due to potential contamination from the oropharynx [22]. The collection of induced sputum is noninvasive; however, it requires preliminary inhalation of an aerosol and is often impractical. The EBC collection technique is a safe that can be repeated easily, and the sample contains nonvolatile water-soluble compounds from the airway mucosa, which are likely to reflect the composition of the airway lining fluid[23,24]. Therefore, we used EBC-derived EVs to identify the bacterial composition of patients with asthma for the first time.

The relationship between asthma and lung bacterial community diversity is controversial. Previous studies in patients with asthma have documented increased bacterial burden and diversity using induced sputum and bronchial brushing compared with healthy controls[2,25]. Similarly, other studies reported an increase in the bacterial diversity in BAL samples obtained from 23 patients with eosinophil-low asthma [26]. This study also showed greater bacterial diversity in patients with asthma based on Chao1, Shannon, and Simpson indices compared with controls using EVs derived from EBC. Conversely, other studies failed to replicate these findings and occasionally presented opposite results. A study reported that there were no significant differences in alpha diversity between patients with severe and non-severe asthmatics, and healthy controls[27,28]. This discrepancy may be attributed to the differences in the populations from which samples were obtained and in the sampling collection methods and sites.

In this study, EBC EV samples obtained from patients with asthma had a significantly higher proportion of *Sphingomonas* ,*Akkermansia* and *Methylophaga* at the genus level compared to the controls. *Sphingomonas* species have been associated with increased airway hyperresponsiveness to methacholine and eosinophil-high asthma [2,26]. Firmicutes and Proteobacteria were phyla that were demonstrated to be increased in the patients with asthma in the GBM ad ANN asthma models. A previous study reported that relevant differences were found in Firmicutes and Proteobacteria between induced sputum samples obtained from asthmatics and healthy controls[25], which is consistent with our results. Other studies using bronchial samples reported that Proteobacteria were present in higher proportions in patients with asthma compared to the controls [2,3]. Elevated levels of the Firmicutes phylum have been found in the lung tissue-derived EVs of patients with COPD, which may be explained by differences in pH, oxygenation levels, and temperature of the lung environment related to esophageal reflux[18]. Patients with chronic lung diseases such as asthma and COPD share similar bacteria causing dysbiosis, such as Protebacteria and Firmicutes[29]. Although we can speculate on the causal relationship between the microbiota and airway inflammation, we cannot confirm its existence yet. Further studies are needed to investigate the specific causal relationship between the microbiota and airway inflammation mediated via the molecular pathways. Additionally, based on the result

of the microbiome-based differences between patients with asthma and controls, we proposed a diagnostic AI model of asthma with good performance. This might be a biomarker for risk assessment and treatment of asthma through further studies.

As the scope of microbiome-focused research widens, numerous studies have reported distinct respiratory microbiota according to the various asthma phenotypes and endotypes, such as obesity, severity, airway hyperresponsiveness, asthma control, acute exacerbation, and response to steroids [30-34]. In particular, the eosinophilic inflammatory phenotype and underlying T2-high endotype corresponded to lung microbial communities with comparatively lower bacterial loads and substantial diversity [35]. Unfortunately, we did not demonstrate the distinction in the airway microbiota according to blood eosinophil counts in this study. Our results may be attributed to the relatively small number of EBC samples and variations in the participants, which may be potentially attributable to populations, environment, and antibiotic use history.

Recent immunomodulatory treatments for asthma mainly focused on the mechanisms of type 2 inflammation, such as biologics, i.e., antibodies against cytokine mediators (e.g., IL-5, IL-4, and IL-13)[31]. However, these treatments are not effective for patients with a T2-low endotype, which may limit the indications of biologics in them. The strong association between airway dysbiosis and pathogenesis of neutrophilic asthma has also been reported[36,37]; thus, the manipulation of the airway and gut microbiomes may be a novel therapeutic strategy for asthma, especially in patients with T2-low endotypes. Further investigation into the impact of microbiota therapies may be a useful tool for controlling asthma using a personalized medicine approaches.

This study has several limitations. Firstly, the sample size of both groups was relatively small. Larger studies are needed to confirm our finding. Second, baseline characteristics such as dietary patterns and use of antibiotics were not evaluated in controls and asthmatics, which might have affected the bacterial community as a confounding bias. Nevertheless, to the best our knowledge, this was the first study to analyze the microbiota using EBC-derived EVs in patients with asthma and compare it with healthy controls.

In conclusion, we have demonstrated differences between the lung microbiome of patients with asthma and controls by analyzing EVs derived from EBC. We also proposed an asthma diagnostic model using AI. Our findings suggest that asthma may be associated with dysbioisis of the bacterial community, and the microbiome obtained using EBC EVs may be a potential indicator for the diagnosis of asthma. Further studies are needed to verify the distinct patterns of the microbiome according to asthma phenotypes and endotypes, considering the heterogeneity of asthma.

References

1. Pavord ID, Beasley R, Agusti A, et al. After asthma: redefining airways diseases. *Lancet.* 2018;391:350-400.

2. Huang YJ, Nelson CE, Brodie EL, et al. Airway microbiota and bronchial hyperresponsiveness in patients with suboptimally controlled asthma. *J Allergy Clin Immunol.* 2011;127:372-381 e371-373.

3. Hilty M, Burke C, Pedro H, et al. Disordered microbial communities in asthmatic airways. *PLoS One.* 2010;5:e8578.

4. Sharma A, Laxman B, Naureckas ET, et al. Associations between fungal and bacterial microbiota of airways and asthma endotypes. *J Allergy Clin Immun.* 2019;144:1214-+.

5. Taylor SL, Leong LEX, Choo JM, et al. Inflammatory phenotypes in patients with severe asthma are associated with distinct airway microbiology. *J Allergy Clin Immun.* 2018;141:94-+.

6. Lee JH, Choi JP, Yang J, et al. Metagenome analysis using serum extracellular vesicles identified distinct microbiota in asthmatics.*Sci Rep-Uk.* 2020;10.

7. Pyun BY. Extracellular Vesicle: An Unknown Environmental Factor for Causing Airway Disease. *Allergy Asthma Immun.* 2016;8:179-180.

8. EL Andaloussi S, Maeger I, Breakefield XO, Wood MJA. Extracellular vesicles: biology and emerging therapeutic opportunities. *Nat Rev Drug Discov.* 2013;12:348-358.

9. Kaparakis-Liaskos M, Ferrero RL. Immune modulation by bacterial outer membrane vesicles. *Nat Rev Immunol.* 2015;15:375-387.

10. Hunt J. Exhaled breath condensate: An evolving tool for noninvasive evaluation of lung disease. *J Allergy Clin Immun.* 2002;110:28-34.

11. An J, Lee JH, Sim JH, et al. Serum Eosinophil-Derived Neurotoxin Better Reflect Asthma Control Status Than Blood Eosinophil Counts.*J Allergy Clin Immunol Pract.* 2020;8:2681-2688 e2681.

12. Kim TB, Jang AS, Kwon HS, et al. Identification of asthma clusters in two independent Korean adult asthma cohorts. *Eur Respir J.*2013;41:1308-1314.

13. Yang J, McDowell A, Kim EK, et al. Consumption of a Leuconostoc holzapfelii-enriched synbiotic beverage alters the composition of the microbiota and microbial extracellular vesicles. *Exp Mol Med.*2019;51.

14. Yang J, Moon HE, Park HW, et al. Brain tumor diagnostic model and dietary effect based on extracellular vesicle microbiome data in serum.*Exp Mol Med.* 2020;52:1602-1613.

15. Abadi M, Agarwal A, Barham P, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:160304467.* 2016.

16. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research.*2011;12:2825-2830.

17. Dorward DW, Schwan T, Garon CF. Immune capture and detection of Borrelia burgdorferi antigens in urine, blood, or tissues from infected ticks, mice, dogs, and humans. *J Clin Microbiol.*1991;29:1162-1170.

18. Kim HJ, Kim YS, Kim KH, et al. The microbiome of the lung and its extracellular vesicles in nonsmokers, healthy smokers and COPD patients.*Exp Mol Med.* 2017;49.

19. Samra M, Nam SK, Lim DH, et al. Urine Bacteria-Derived Extracellular Vesicles and Allergic Airway Diseases in Children. *Int Arch Allergy Imm.* 2019;178:150-158.

20. Yang J, McDowell A, Seo H, et al. Diagnostic Models for Atopic Dermatitis Based on Serum Microbial Extracellular Vesicle Metagenomic Analysis: A Pilot Study. *Allergy Asthma Immun.* 2020;12:792-805.

21. Choi EB, Hong SW, Kim DK, et al. Decreased diversity of nasal microbiota and their secreted extracellular vesicles in patients with chronic rhinosinusitis based on a metagenomic analysis. *Allergy.*2014;69:517-526.

22. Sullivan A, Hunt E, MacSharry J, Murphy DM. The Microbiome and the Pathophysiology of Asthma. *Resp Res.* 2016;17.

23. Montuschi P, Barnes PJ. Analysis of exhaled breath condensate for monitoring airway inflammation. *Trends Pharmacol Sci.*2002;23:232-237.

24. Cherot-Kornobis N, Hulo S, de Broucker V, et al. Induced Sputum, Exhaled NO, and Breath Condensate in Occupational Medicine. *J Occup Environ Med.* 2012;54:922-927.

25. Marri PR, Stern DA, Wright AL, Billheimer D, Martinez FD. Asthma-associated differences in microbial composition of induced sputum. *J Allergy Clin Immunol.* 2013;131:346-352 e341-343.

26. Sverrild A, Kiilerich P, Brejnrod A, et al. Eosinophilic airway inflammation in asthmatic patients is associated with an altered airway microbiome. *J Allergy Clin Immun.* 2017;140:407-+.

27. Zhang QL, Cox M, Liang ZK, et al. Airway Microbiota in Severe Asthma and Relationship to Asthma Severity and Phenotypes. *Plos One.*2016;11.

28. Li NJ, Qiu RH, Yang ZW, et al. Sputum microbiota in severe asthma patients: Relationship to eosinophilic inflammation. *Resp Med.*2017;131:192-198.

29. Marsland BJ, Gollwitzer ES. Host-microorganism interactions in lung diseases. *Nat Rev Immunol.* 2014;14:827-835.

30. Frati F, Salvatori C, Incorvaia C, et al. The Role of the Microbiome in Asthma: The Gut-Lung Axis. *Int J Mol Sci.* 2019;20.

31. Kozik AJ, Huang YJ. The microbiome in asthma Role in pathogenesis, phenotype, and response to treatment. *Ann Allerg Asthma Im.*2019;122:270-275.

32. Qiu RH, Li NJ, Yang ZW, et al. Analysis of the Sputum Microbiome in the Severe Asthma. *Chest.* 2016;149:14a-14a.

33. Goleva E, Jackson LP, Harris JK, et al. The Effects of Airway Microbiome on Corticosteroid Responsiveness in Asthma. *Am J Resp Crit Care.* 2013;188:1193-1201.

34. Gilstrap DL, Kraft M. Asthma and the host-microbe interaction.*J Allergy Clin Immun.* 2013;131:1449-U1338.

35. Loverdos K, Bellos G, Kokolatou L, et al. Lung Microbiome in Asthma: Current Perspectives. *J Clin Med.* 2019;8.

36. Thomson NC. Novel approaches to the management of noneosinophilic asthma. *Ther Adv Respir Dis.* 2016;10:211-234.

37. Yang X, Li HN, Ma QL, Zhang Q, Wang CZ. Neutrophilic Asthma Is Associated with Increased Airway Bacterial Burden and Disordered Community Composition. *Biomed Res Int.* 2018.

Table 1. Basic clinical characteristics of the study population

Group	Institution	No.	Sex (F/M)	Age [mean (SD)]
Controls	Asan Medical Center	58	1.6	50.8 (SD: 14.4)
Asthmatics	Asan Medical Center	251	2.4	32.4 (SD: 7.2)

SD: Standard Deviation.

Table 2. AUC performance of EBC asthma diagnostic models

AUC	Logistic (T test)	Logistic (LEfSe)	GBM	ANN	Ensemble
Mean	0.7531	0.7440	0.8325	0.7686	0.8257
Median	0.7485	0.7569	0.8311	0.7719	0.8165
SD	0.0287	0.0423	0.0495	0.0293	0.0337
Maximum	0.7939	0.7968	0.9547	0.8077	0.9020
Minimum	0.7025	0.6791	0.7705	0.7032	0.7785

AUC, area under the curve; EBC, exhaled breath condensate; LefSe, linear discriminant analysis effect size; GBM, gradient boosting machine; ANN, artificial neural network; SD, standard deviation

Table 3. Common important features between the top 50 ranked features in GBM and ANN asthma models

No.	**Phylum**	**Class**	**Order**	**Family**	**Genus**
1	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae UCG-01
2	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Lachnospiraceae UCG-008
3	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas
4	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter

No.	Phylum	Class	Order	Family	Genus
5	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	[Eubacterium] hallii group
6	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Blautia
7	Actinobacteria	Actinobacteria	Bifidobacteriales	Bifidobacteriaceae	Bifidobacterium
8	Actinobacteria	Coriobacteriia	Coriobacteriales	Coriobacteriaceae	Collinsella
9	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Paracoccus
10	Firmicutes	Erysipelotrichia	Erysipelotrichales	Erysipelotrichaceae	Holdemanella
11	Actinobacteria	Actinobacteria	Corynebacteriales	Corynebacteriaceae	Lawsonella
12	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	Enterococcus
13	Actinobacteria	Actinobacteria	Propionibacteriales	Propionibacteriaceae	Cutibacterium
14	Actinobacteria	Actinobacteria	Micrococcales	Intrasporangiaceae	Intrasporangium
15	Proteobacteria	Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	Acidovorax
16	Firmicutes	Erysipelotrichia	Erysipelotrichales	Erysipelotrichaceae	Erysipelotrichaceae UCG-0
17	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Dorea
18	Firmicutes	Clostridia	Clostridiales	Clostridiaceae 1	Clostridium sensu stricto 1
19	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Alcanivoracaceae	Alcanivorax
20	Proteobacteria	Alphaproteobacteria	Thalassobaculales	Thalassobaculaceae	Thalassobaculum
21	Proteobacteria	Gammaproteobacteria	Nitrosococcales	Methylophagaceae	Methylophaga

GBM: gradient boosting machine; ANN: artificial neural network.

Figure legends

Figure 1. Alpha and beta diversity of the EV-extracted microbiome obtained from EBC samples

A) Alpha diversity within each sample of the control and asthma EBC samples was measured based on the observed, Chao1, Shannon, and Simpson diversity indices and plotted using a boxplot for each clinical group. B) Bray-Curtis distance between the samples was measured and visualized as a PCoA plot for the dissimilarity between samples.

Figure 2. Genus-level microbial EV compositional differences between the clinical samples

A) The average genus-level composition of the microbial EVs in the control and asthma groups. B) Heat-map depicting the differential relative abundance values of the genera in each sample in the control and asthma group. C) Average relative abundance of genera in control and asthma samples with significant differences (*P <0.05 and **P<0.01).

Figure 3. Asthma diagnostic model performance evaluation

A) Box plot of the AUC values of 10 iterations of each EBC asthma diagnostic model method: Logistic (T test), Logistic (LEfSe), GBM, ANN, and ensemble. B) ROC curve plotting the AUC values of 10 iterations of each method against the specificity and sensitivity values of each model.

Figure 1.

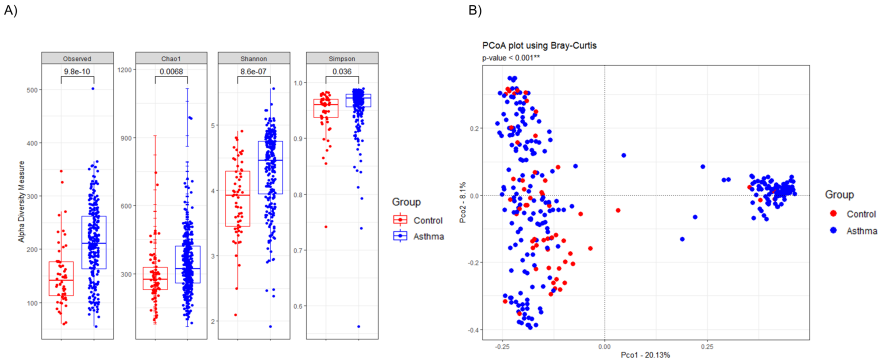

Figure 2.

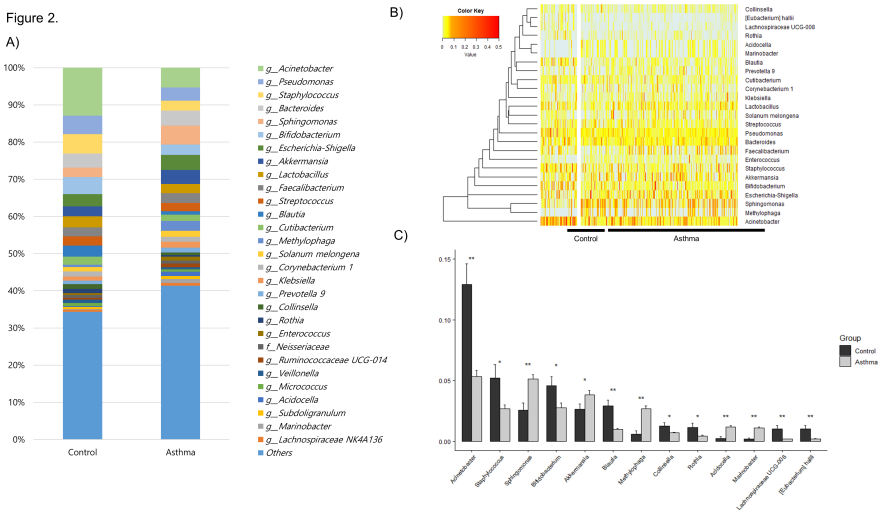

Figure 3.