Asthma and lower airway disease

Inflammatory phenotypes in patients with severe asthma are associated with distinct airway microbiology

Steven L. Taylor, BSc,a,b* Lex E. X. Leong, PhD,a,b* Jocelyn M. Choo, PhD,a,b Steve Wesselingh, FRACP, PhD,a,b Ian A. Yang, FRACP, PhD,c,d John W. Upham, FRACP, PhD,c,d* Paul N. Reynolds, MBBS, PhD,d,f,g Sandra Hodge, PhD,f,g Alan L. James, FRACP, PhD,b,l Christine Jenkins, MBBS, FRACP j,k Matthew J. Peters, MD, FRACP,k,l Melissa Baraket, FRACP, PhD,m,p Guy B. Marks, MBBS, PhD,m,n,p Peter G. Gibson, MBBS, FRACP,n,o Jodie L. Simpson, PhD,o and Geraint B. Rogers, PhD,a,b,l Adelaide, St Lucia, Chermside, Wooloongabba, Nedlands, Crawley, Newtown, North Ryde, Concord, Liverpool, Glebe, Callaghan, and Sydney, Australia

From a the South Australian Health and Medical Research Institute, Adelaide; b the SAHMRI Microbiome Research Laboratory, School of Medicine, Flinders University, Adelaide; c the School of Medicine, University of Queensland, St Lucia; d the Department of Thoracic Medicine, The Prince Charles Hospital, Chermside; e the Translational Research Institute, Princess Alexandra Hospital, Woolloongabba; f the Department of Thoracic Medicine, Royal Adelaide Hospital and Lung Research Laboratory, Hanson Institute, Adelaide; g the School of Medicine, University of Adelaide, Adelaide; h the Department of Pulmonary Physiology and Sleep Medicine, Sir Charles Gardiner Hospital, Nedlands; i the School of Medicine and Pharmacology, University of Western Australia, Crawley; j Respiratory Trials, George Institute for Global Health, Newtown; k the Australian School of Advanced Medicine, Macquarie University, North Ryde; l the Department of Thoracic Medicine, Concord General Hospital; m the Respiratory Medicine Department and Ingham Institute, Liverpool Hospital; n the Woolcock Institute of Medical Research, Glebe; o Respiratory and Sleep Medicine, Priority Research Centre for Healthy Lungs, University of Newcastle, Callaghan; and p the South Western Sydney Clinical School, University of New South Wales, Sydney.

*These authors contributed equally to this work.
†These authors contributed equally to this work.

Supported by the Australian National Health and Medical Research Council (NHMRC); grant no. 569246 and a grant from the John Hunter Hospital Charitable Trust.

Disclosure of potential conflict of interest: I. A. Yang’s, J. W. Upham’s, and C. Jenkins’ institution received a grant from National Health & Medical Research Council (NHMRC), Australia, for this work. P. N. Reynolds’ and J. L. Simpson’s institution received a grant from John Hunter Charitable Trust and the NHMRC CRE Severe Asthma for this work. S. Hodge is employed by the University of Adelaide; her institution has grants with NHMRC; and she received royalties from the book Lung Macrophages in Health and Disease. O. B. Marks’ institution received grants from AstraZeneca and GlaxoSmithKline for other works. P. G. Gibson’s institution received a grant from NHMRC for this work and has grants from the NHMRC, AstraZeneca, and GlaxoSmithKline for other works; he has personally received payment for lectures from AstraZeneca, GlaxoSmithKline, and Novartis. The rest of the other authors declare that they have no relevant conflicts of interest.

Received for publication October 16, 2016; revised February 28, 2017; accepted for publication March 15, 2017.

Available online May 4, 2017.

Corresponding author: Geraint B. Rogers, PhD, School of Medicine, Flinders University, University Drive, Bedford Park, Adelaide SA 5042, Australia. E-mail: geraint.rogers@sahmri.com.

The CrossMark symbol notifies online readers when updates have been made to the article such as errata or minor corrections 0091-6749 © 2017 The Authors. Published by Elsevier Inc. on behalf of the American Academy of Allergy, Asthma & Immunology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

http://dx.doi.org/10.1016/j.jaci.2017.03.044
Background: Asthma pathophysiology and treatment responsiveness are predicted by inflammatory phenotype. However, the relationship between airway microbiology and asthma phenotype is poorly understood.

Objective: We aimed to characterize the airway microbiota in patients with symptomatic stable asthma and relate composition to airway inflammatory phenotype and other phenotypic characteristics.

Methods: The microbial composition of induced sputum specimens collected from adult patients screened for a multicenter randomized controlled trial was determined by using 16S rRNA gene sequencing. Inflammatory phenotypes were defined by sputum neutrophil and eosinophil cell proportions. Microbiota were defined by using β- and α-diversity measures, and interphenotype differences were identified by using similarity of percentages, network analysis, and taxon fold change. Phenotypic predictors of airway microbiology were identified by using multivariate linear regression.

Results: Microbiota composition was determined in 167 participants and classified as eosinophilic (n = 84), neutrophilic (n = 14), paucigranulocytic (n = 60), or mixed neutrophilic-eosinophilic (n = 9) asthma phenotypes. Airway microbiology was significantly less diverse (P = .022) and more dissimilar (P = .005) in neutrophilic compared with eosinophilic participants. Sputum neutrophil proportions, but not eosinophil proportions, correlated significantly with these diversity measures (β-diversity: Spearman r = –0.374, P < .001; α-diversity: r = 0.238, P = .002). Interphenotype differences were characterized by a greater frequency of pathogenic taxa at high relative abundance and reduced Streptococcus, Gemella, and Porphyromonas taxa relative abundance in patients with neutrophilic asthma. Multivariate regression confirmed that sputum neutrophil proportion was the strongest predictor of microbiota composition.

Conclusions: Neutrophilic asthma is associated with airway microbiology that is significantly different from that seen in patients with other inflammatory phenotypes, particularly eosinophilic asthma. Differences in microbiota composition might influence the response to antimicrobial and steroid therapies and the risk of lung infection. (J Allergy Clin Immunol 2018;141:94-103.)

Key words: Asthma, microbiome, neutrophil, eosinophil

Evidence-based patient stratification is required to address the heterogeneity of asthma severity and treatment responsiveness. Asthma phenotype based on characteristics of airway inflammation is increasingly recognized as an important prognostic indicator. In addition to an allergen-induced, Th2 lymphocyte, IL-5–mediated, eosinophilic inflammatory response, asthma can also occur in the absence of eosinophilic inflammation (termed noneosinophilic asthma). Indeed, based on relative numbers of sputum eosinophils and neutrophils, 4 inflammatory subtypes have been described: eosinophilic asthma, neutrophilic asthma, mixed granulocytic asthma, and paucigranulocytic asthma.

Unlike the relatively well-defined mechanisms that result in eosinophilic airway inflammation, those leading to noneosinophilic asthma, particularly neutrophilic asthma, remain relatively poorly understood. Furthermore, although noneosinophilic phenotypes occur across the spectrum of asthma severity, they typically respond poorly to corticosteroids. Inflammatory phenotypes have also been shown to differ with respect to airway microbiology. Compared with other patients with asthma, those with neutrophilic asthma are more likely to have a potentially pathogenic organism identified by means of either culture-based or culture-independent approaches and have reduced airway bacterial diversity. Given that airway microbiota composition is associated with the degree of airway hyperresponsiveness among patients with suboptimally controlled asthma, interphenotype differences in airway microbiology are likely to be clinically important.

The relationships between asthma inflammatory phenotypes and airway microbiology are likely to be complex and bidirectional. Asthma phenotypes represent immunologic and physicochemical differences within the lower airways that are likely to be reflected through their selective effect on microbial growth and airway clearance in divergent lower airway microbiota. Where these differences involve the increased abundance of particular respiratory pathogens or a depletion of commensal populations, they could contribute substantially to the course of airway disease or risk of adverse treatment events, such as corticosteroid-associated pneumonia. On the other hand, the characteristics of airway microbiology, even in the absence of frank infection, could influence the asthma inflammatory phenotype. Defining the relationships between inflammatory phenotype and lower airway microbiota would inform our understanding of asthma pathophysiology and could help identify prognostic markers.

Several previous studies have reported differences in airway microbiology in patients with eosinophilic and noneosinophilic asthma and the existence of significant relationships between this microbiota composition and clinical asthma measures. However, although providing important insight, these studies have involved relatively small and heterogeneous patient cohorts. Our study, based on participants enrolled in the Asthma and Macrolides: the Azithromycin Efficacy and Safety Study (ACTRN12609000197235), was more than 3 times the size of any study performed previously and focused on a well-defined population of patients with severe but stable asthma, the majority of whom were treated with inhaled corticosteroids (ICSs).

Through application of a systematic approach to microbiota characterization, we aimed to assess whether asthma inflammatory phenotypes were associated with substantially different lower airway bacteriology (herein referred to as microbiology), to identify bacterial taxa that discriminate among inflammatory phenotypes, and to determine the contribution of patient and clinical characteristics to variation in the composition of the bacterial component of the microbiome.

METHODS

Study population

For full methodology, see the Methods section in this article’s Online Repository at www.jacionline.org. Analysis was performed on samples collected as part of the baseline screening population from the clinical Azithromycin Efficacy and Safety Study (ACTRN12609000197235). Participants were recruited...
from 8 centers across Australia (detailed in the Methods section in this article’s Online Repository). Predefined inflammatory phenotype categories based on sputum cell counts relative to patient age were assigned, as published previously and detailed in the Methods section in this article’s Online Repository. Briefly, the neutrophilic phenotype was defined as 61% or greater neutrophils (neutrophil percentage cutoff dependent on age), the eosinophilic phenotype as 3% or greater eosinophils, the paucigranulocytic phenotype as 61% or less neutrophils and 3% or less eosinophils, and the mixed granulocytic phenotype as 61% or greater neutrophils and 3% or greater eosinophils.

DNA extraction, 16S rRNA gene amplicon sequencing, and gene copy numbers

DNA extraction was performed on 100-μL sputum aliquots by using a combined physical, enzymatic, and heat-based cell lysis, followed by phenol chloroform extraction and DNA recovery with EZ-10 Spin columns (Bio Basic, Ontario, Canada). The V1-3 hypervariable region of the bacterial 16S rRNA gene was amplified from sputum DNA by using the modified primers 27F and 519R. Amplicons were cleaned, indexed, and sequenced according to the Illumina MiSeq 16S Metagenomic Sequencing Library Preparation Protocol (Illumina, San Diego, Calif). 16S rRNA sequence data were processed, as previously described. Spurious operational taxonomic units (OTUs) were removed systematically by using previous reports of common laboratory sequencing contaminants. Bacterial burden was quantified with quantitative PCR (qPCR) for the 16S rRNA gene. Detailed extraction, sequencing, and qPCR protocols are described in the Methods section in this article’s Online Repository.

Diversity measurements and statistical analyses

Five α-diversity (within-sample variance) indices were used to test a variety of parameters of within-patient taxon distribution: Faith’s phylogenetic diversity (in which a higher value indicates a more phylogenetically diverse sample), Simpson’s and Pielou’s evenness indices (in which a higher value indicates a more equitable distribution of taxa abundance), taxa richness (the total number of taxa detected), and Shannon-Weiner diversity (a measure incorporating both the number and abundance), taxa richness (the total number of taxa detected), and Shannon-Weiner index, Simpson index, and Pielou evenness; see Figs E2 and E3 in this article’s Online Repository at www.jacionline.org).

RESULTS

Clinical characteristics

Induced sputum samples were obtained from 187 participants. Of these, 13 were excluded because of poor sample quality. Of the 174 that underwent 16S rRNA gene amplicon sequencing, a further 7 were excluded because of an insufficient sequence read depth (see Table E1 in this article’s Online Repository at www.jacionline.org). The remaining 167 subjects were classified as one of 4 inflammatory phenotypes based on previously described sputum inflammatory cell count percentages: neutrophilic (n = 14), eosinophilic (n = 84), paucigranulocytic (n = 60), or mixed granulocytic (n = 9). There was no significant difference in age, sex distribution, atopy, smoking history, ICS dose, Global Initiative for Asthma treatment step, or mean Asthma Control Questionnaire score between these phenotypic groups, as assessed by using multiple comparison tests (Table I). However, there were significant differences in lung function, as assessed by both FEV₁ percent predicted (P = .035) and FEV₁/forced vital capacity percentage (P = .013).

After quality filtering and chimera removal, 16S rRNA gene amplicon sequencing resulted in a median read depth of 12,792 (quartile 1 and quartile 3, 8,060 and 16,595). Sequence data were subsampled to a uniform depth of 1,732 reads based on rarefaction curve asymptotes and Good’s coverage values. No significant differences in total bacterial burden were found between inflammatory phenotypes (P = .51, Kruskal-Wallis test; see Fig E1 in this article’s Online Repository at www.jacionline.org).

α-Diversity

Participants with neutrophilic asthma had significantly lower Faith’s phylogenetic scores (P = .022) than participants with eosinophilic asthma, which resembled those of patients with paucigranulocytic asthma (Fig 1, A). Faith’s phylogenetic diversity significantly correlated with the sputum neutrophil percentage (r = −.0374, P < .0001; Fig 1, B) but not with the sputum eosinophil percentage (r = .0146, P = .060; Fig 1, C). Analysis with a range of alternative α-diversity indices (taxa richness, Shannon-Wiener index, Simpson index, and Pielou evenness; see Figs E2 and E3 in this article’s Online Repository at www.jacionline.org) resulted in consistent findings in relation to phenotype, sputum neutrophil percentage, and sputum eosinophil percentage. Together, these results demonstrate a significant relationship between airway microbiota composition and sputum neutrophilia but not sputum eosinophilia.

β-Diversity

Principal coordinate analysis of weighted UniFrac similarity distance showed that neutrophilic samples were distinguished from other phenotypes along the first and second principal coordinates, whereas the other phenotypes broadly clustered together (Fig 2, A). Consistent with these observations, a permutational multiple ANOVA (PERMANOVA) test showed that phenotype grouping contributed significantly to differences in microbial composition of the samples (P = .0004, pseudo-F = 3.997; see Table E2 in this article’s Online Repository at www.jacionline.org). Pairwise PERMANOVA comparing the phenotype groups indicated that variance was attributed to the neutrophilic versus eosinophilic (P = .0001, T = 3.30) and neutrophilic versus paucigranulocytic (P = .0015, T = 2.52)
TABLE I. Clinical and inflammatory cell parameters of participants

<table>
<thead>
<tr>
<th></th>
<th>Neutrophil</th>
<th>Eosinophil</th>
<th>Paucigranulocytic</th>
<th>Mixed granulocytic</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>14</td>
<td>84</td>
<td>60</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Age (y), mean (SD)</td>
<td>59.8 (13.9)</td>
<td>57.2 (15.2)</td>
<td>55.8 (13.8)</td>
<td>56.8 (17.8)</td>
<td>.817</td>
</tr>
<tr>
<td>Male sex, no. (%)</td>
<td>9 (64.3)</td>
<td>30 (35.7)</td>
<td>23 (38.3)</td>
<td>6 (66.7)</td>
<td>.084</td>
</tr>
<tr>
<td>Atopic, no. (%)</td>
<td>11 (78.6)</td>
<td>62 (76.5)</td>
<td>52 (86.7)</td>
<td>6 (66.7)</td>
<td>.305</td>
</tr>
<tr>
<td>Previous smoker, no. (%)</td>
<td>6 (42.9)</td>
<td>32 (38.1)</td>
<td>17 (28.3)</td>
<td>4 (44.4)</td>
<td>.491</td>
</tr>
<tr>
<td>Smoking pack years, median (Q1, Q3)</td>
<td>22.0 (18.6, 27.5)</td>
<td>5.3 (1.2, 15.9)</td>
<td>5.0 (1.3, 30.0)</td>
<td>4.4 (0.9, 63.6)</td>
<td>.379</td>
</tr>
<tr>
<td>Duration of asthma (y), median (Q1, Q3)</td>
<td>33.8 (3.5, 48.3)</td>
<td>36.3 (19.7, 49.2)</td>
<td>33.7 (14.2, 54.4)</td>
<td>53.6 (32.6, 60.4)</td>
<td>.205</td>
</tr>
<tr>
<td>FEV₁ (% predicted), mean (SD)</td>
<td>70.3 (18.2)</td>
<td>70.0 (17.9)</td>
<td>78.7 (19.2) *</td>
<td>68.4 (17.3)</td>
<td>.035</td>
</tr>
<tr>
<td>FVC (% predicted), mean (SD)</td>
<td>81.4 (11.6)</td>
<td>83.6 (16.4)</td>
<td>85.1 (15.0)</td>
<td>82.9 (11.7)</td>
<td>.854</td>
</tr>
<tr>
<td>FEV₁/FVC (%), mean (SD)</td>
<td>65.8 (14.5)</td>
<td>64.9 (12.2)</td>
<td>71.3 (12.2) *</td>
<td>63.9 (12.7)</td>
<td>.013</td>
</tr>
<tr>
<td>ACQ6 score, mean (SD)</td>
<td>2.1 (1.2)</td>
<td>1.9 (0.9)</td>
<td>1.6 (0.8)</td>
<td>1.5 (0.6)</td>
<td>.210</td>
</tr>
<tr>
<td>GINA treatment step</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.058</td>
</tr>
<tr>
<td>1</td>
<td>1 (7.1)</td>
<td>2 (2.5)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0 (0)</td>
<td>10 (12.4)</td>
<td>12 (20.3)</td>
<td>2 (22.2)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>11 (78.6)</td>
<td>67 (82.7)</td>
<td>47 (80.0)</td>
<td>7 (77.8)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2 (14.3)</td>
<td>2 (2.5)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>ICS dose (µg), median (Q1, Q3)</td>
<td>9.8 (7.6, 12.7) *</td>
<td>3.6 (1.9, 7.6)</td>
<td>3.74 (2.0, 7.6)</td>
<td>8.64 (5.22, 11.34)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Total cell count (× 10⁵/mL, [Q1, Q3])</td>
<td>2000 (1280, 2000), n = 13</td>
<td>1000 (800, 2000), n = 82</td>
<td>1000 (800, 2000), n = 59</td>
<td>1600 (1000, 2000)</td>
<td>.254</td>
</tr>
<tr>
<td>Viability (%)</td>
<td>89.2 (73.5, 93.0) *</td>
<td>60.8 (52.2, 80.7)</td>
<td>68.9 (55.7, 79.5)</td>
<td>90.0 (84.4, 93.7) *</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Neutrophils (%)</td>
<td>75.0 (68.80, 84.00) *</td>
<td>27.13 (14.38, 41.00)</td>
<td>34.13 (12.63, 49.25)</td>
<td>77.00 (71.25, 77.50) *</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Eosinophils (%)</td>
<td>0.63 (0.50, 1.00) *</td>
<td>6.88 (3.63, 18.13)</td>
<td>0.25 (0.00, 1.00) *</td>
<td>6.75 (3.25, 12.25) *</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Macrophages (%)</td>
<td>20.75 (14.50, 30.25) *</td>
<td>52.75 (38.25, 71.00)</td>
<td>53.34 (43.50, 75.63)</td>
<td>16.00 (15.50, 20.25) *</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Lymphocytes (%)</td>
<td>0.63 (0.25, 1.00) *</td>
<td>1.13 (0.25, 2.25)</td>
<td>0.88 (0.25, 2.63)</td>
<td>0.25 (0.00, 0.38) *</td>
<td>&lt;.018</td>
</tr>
<tr>
<td>Columnar epithelial cells (%)</td>
<td>0.25 (0.00, 1.25) *</td>
<td>2.17 (0.75, 6.00)</td>
<td>4.13 (1.21, 8.50)</td>
<td>0.50 (0.25, 1.50) *</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Squamous cells (%)</td>
<td>1.11 (0.50, 3.85) *</td>
<td>5.99 (2.32, 16.23)</td>
<td>6.18 (2.92, 11.31)</td>
<td>3.38 (0.25, 8.05) *</td>
<td>.002</td>
</tr>
</tbody>
</table>

P values in the last column describe variance across 4 phenotypes. ACQ6, Asthma Control Questionnaire 6; FVC, forced vital capacity; GINA, Global Initiative for Asthma; Q1, Q3, quartile 1, quartile 3. *P < .05 versus eosinophilic asthma. †P < .05 versus paucigranulocytic asthma. ‡P < .05 versus neutrophilic asthma.

groups (see Table E3 in this article’s Online Repository at www.jacionline.org).

Assessment of microbiota dispersion based on distance from centroid was consistent with PERMANOVA analysis, with the samples from neutrophilic phenotype participants having significantly higher distances from the centroid than samples from eosinophilic and paucigranulocytic participants (Fig 2, B). In keeping with α-diversity analyses, variance in distance from centroid was associated with sputum neutrophil percentage rather than sputum eosinophil percentage (see Fig E4 in this article’s Online Repository at www.jacionline.org). β-Diversity analyses using a second distance measure, Bray-Curtis similarity, produced consistent findings (see Fig E5 in this article’s Online Repository at www.jacionline.org).

Taxon distribution and network analysis

SIMPER analysis was used to rank taxa according to their contribution to intergroup variance in microbiota composition. Thirteen taxa were identified, which cumulatively accounted for approximately 50% of total variance between neutrophilic and eosinophilic samples (see Table E4 in this article’s Online Repository at www.jacionline.org). Hierarchical cluster analysis based on relative taxon abundance revealed that Moraxella and Haemophilus clustered separately from the other 11 taxa (Fig 3). In the patients with neutrophilic asthma, Moraxella and Haemophilus taxa exceeded 40% relative abundance in 6 (42.9%) of 14 samples compared with only 1 (1.19%) of 84 patients with eosinophilic asthma, 7 (11.7%) of 60 patients with paucigranulocytic asthma, and 1 (11.1%) of 9 patients with a mixed phenotype (χ² = 25.5, P < .0001; Fig 3). Relationships between bacterial taxon relative abundance were further visualized by using network analysis (Fig 4 and see Fig E6 in this article’s Online Repository at www.jacionline.org), revealing a bacterial community of taxa with positively correlated abundances in almost all cases. Most of these taxa were more prevalent in eosinophilic samples than in neutrophilic samples. Haemophilus taxon, which had a mean abundance that was higher in neutrophilic samples, was the single exception, negatively correlating with other members of the sputum bacterial community.

Of the 13 discriminant taxa identified by SIMPER, Streptococcus II (see the Methods section in this article’s Online Repository at www.jacionline.org for classification), Gemella, Rothia, and Porphyromonas taxa were significantly less abundant in neutrophilic than in eosinophilic and paucigranulocytic phenotypes (Fig 5, A). Sputum neutrophil percentage positively correlated with the relative abundance of Moraxella taxon and negatively correlated with the relative abundance of Streptococcus I, Gemella, and Porphyromonas taxa (Fig 5, B). In contrast, Haemophilus taxon negatively correlated with eosinophil percentage, and Streptococcus I, Neisseria, and Gemella...
taxa positively correlated with eosinophil percentage (see Fig E7 in this article’s Online Repository at www.jacionline.org). Prevotella, Actinomyces, Leptotrichia, and Veillonella taxa, although identified by using SIMPER and represented by highly connected nodes in the network analysis, neither differed between airway inflammatory phenotypes nor correlated with sputum cell counts.

Nondominant microbiome
We sought to establish whether differences in microbiota composition between inflammatory phenotypes were explained solely by overgrowth of opportunistic taxa (eg, Haemophilus and Moraxella taxa) or whether differences existed even in the absence of pathogen predominance. Two separate approaches were used to investigate this: rescaling of relative abundance data after exclusion of pathogen predominance and assessment of ranked taxon fold change between neutrophilic and eosinophilic groups based on nonsubsampled taxa counts. Rescaled relative abundance data remained significantly different between inflammatory phenotypes ($P = .0004$, pseudo-$F = 2.38$; see Table E5 in this article’s Online Repository at www.jacionline.org). Pairwise tests revealed significant differences between participants with neutrophilic and eosinophilic phenotypes ($P = .0001$, $T = 2.31$) and between participants with neutrophilic and paucigranulocytic phenotypes ($P = .0002$, $T = 2.26$; see Table E5). Assessment of taxa count—ranked fold change supported these findings, with significant taxa count differences between participants with neutrophilic and eosinophilic phenotypes (see Fig E8 in this article’s Online Repository at www.jacionline.org).

Clinical and inflammatory associations with microbiota composition
In univariate analysis Faith’s phylogenetic diversity significantly inversely correlated with sputum neutrophil percentage, age, and ICS dose and significantly positively correlated with FEV1 percentage (Table II). Conversely, weighted UniFrac distance from centroid significantly positively correlated with sputum neutrophil percentage and was significantly different

FIG. 1. Faith’s phylogenetic diversity is significantly associated with sputum neutrophilia but not eosinophilia. A, Patients grouped by asthma phenotype. B, Neutrophil percentage. The dotted line at 61% neutrophils indicates the phenotype cutoff point. C, Eosinophil percentage. The dotted line at 3% eosinophils indicates the phenotype cutoff point. Colors represent the asthma phenotype: blue is greater than 61% neutrophils, green is greater than 3% eosinophils, yellow is less than 61% neutrophils and less than 3% eosinophils (paucigranulocytic), and purple is both greater than 61% neutrophils and greater than 3% eosinophils (mixed). Statistical significance was assessed by using the Kruskal-Wallis 1-way ANOVA with the Dunn post hoc test (Fig 1, A) or Spearman rank correlation (Fig 1, B and C).

FIG. 2. Microbiota dispersion grouped by asthma phenotype. A, Principal coordinate analysis. The first 2 principal coordinates are plotted on the x- and y-axes, respectively (representing 59.9% of the total variation). B, Distance from centroid. Statistical significance was assessed by using Kruskal-Wallis 1-way ANOVA with the Dunn post hoc test.
based on sex and atopy but not with age, ICS dose, FEV$_1$ percentage, or previous smoking status (Table II). In multivariate analysis, sputum neutrophil percentage was the only variable that independently predicted both Faith’s phylogenetic diversity and weighted UniFrac distance from centroid ($P = .002$ [95% CI, $-0.07$ to $-0.02$] and $P < .001$ [95% CI, 0.07 to 0.22], respectively; Table III). Age and ICS dose both independently predicted Faith’s diversity ($P = .30$ [95% CI = $-0.07$ to $-0.004$] and $P = .42$ [95% CI = $-0.001$ to $-0.001$], respectively), whereas atopy and sex independently predicted distance from centroid ($P = .018$ [95% CI = 1.3 to 8.4] and $P = .039$ [95% CI = $-7.5$ to $-0.30$], respectively).

**DISCUSSION**

To our knowledge, this is the largest study to date to assess predictors of the airway microbiota composition in asthmatic patients. Our primary comparisons were between asthma

---

**FIG 3.** Relative abundance of discriminant taxa among asthma phenotypes. The 13 taxa that collectively contribute to approximately 50% of variance among phenotypes, as determined by using SIMPER analysis. Clustering shows the similarity relationship of genera based on Bray-Curtis similarity distance and the single linkage hierarchical clustering method. *Taxa assigned Actinomyces species uncultured bacteria.

**FIG 4.** Bacterial network analysis of the asthma cohort. Each edge represents a significant correlation colored to indicate either positivity (blue) or negativity (red). Edge width and transparency are proportional to the absolute value of the correlation coefficient. Node size is proportional to mean relative abundance. Node hue is proportional to the difference in taxon relative abundance between the neutrophilic phenotype group and the eosinophilic phenotype group. Correlations were performed with SparCC with a correlation cutoff $R$ value of greater than 0.25 or less than $-0.25$. *Actinomyces species uncultured bacteria. #Actinomyces species oral clone DR002.
FIG 5. Taxa distribution differs by sputum neutrophilia. A, Taxa that significantly differ by patient inflammatory phenotype. B, Significant correlations between taxa and neutrophil percentages. Colors represent asthma phenotype based on neutrophilia or eosinophilia: blue is greater than 61% neutrophils, green is greater than 3% eosinophils, yellow is less than 61% neutrophils and less than 3% eosinophils (paucigranulocytic), and purple is both greater than 61% neutrophils and greater than 3% eosinophils (mixed). The dotted line at 61% neutrophils indicates the phenotype cutoff point. Statistical significance was assessed by using Kruskal-Wallis 1-way ANOVA with the Dunn post hoc test (Fig 5, A) and Spearman rank correlation (Fig 5, B).
inflammatory phenotypes, where we observed significant differences in the composition of airway microbiota. These differences were largely between neutrophilic and eosinophilic phenotypes and reflected a reduced diversity and evenness of detectable bacterial taxa in the neutrophilic participants. Reduced microbiota diversity has been reported after acute and chronic airway infection in neutrophilic patients, reflecting neutrophilic influx into the airways during subclinical lower airway infection, with a reciprocal decrease in the relative abundance of commensal taxa. However, a group of common airway taxa correlated negatively with sputum microbiota percentages (Gemella, Porphyromonas, and Streptococcus taxa) and, importantly, even after controlling for overgrowth effects of Haemophilus and Moraxella taxa, significant differences in microbiota composition between neutrophilic and eosinophilic participants were still observed. This finding suggests that separate phenomena contribute to microbial differences between inflammatory phenotypes: the effect of pathogen overgrowth and the selective pressure of airway inflammatory characteristics in the absence of infection. The latter could result in more broad-scale divergence in composition between the neutrophilic and eosinophilic subgroups, in turn contributing to an increased risk of lower airway infection in neutrophilic patients through an increased presence of opportunistic pathogens.

This finding has clear implications for the clinical management of asthma, in which low-dose macrolide and ICS therapies have been shown to influence overgrowth by opportunistic respiratory pathogens and innate immune function, respectively. Furthermore, the relative lack of efficacy of ICSs in patients with noneosinophilic asthma might lead to use of higher doses compared with those used by eosinophilic patients. The combination of underlying differences in airway microbiota (associated with differences in inflammatory phenotype) and inefficacious therapies being used at higher doses might contribute to reduced bacterial diversity, the high concentrations of Proteobacteria seen in the airways of neutrophilic patients, a greater propensity for lung infection, and a further enhancement of the neutrophilic phenotype.

Although strong associations between the neutrophilic phenotype and sputum microbiota composition were found, associations between eosinophil count and microbiota composition were minimal. This contrasts with previous studies reporting increased Tropheryma taxon associated with eosinophilia and associations between bronchial biopsy eosinophil count and bacterial composition and reduced bacterial burden associated with type 2–high airway inflammation.

An important strength of our study was its involvement of a large group of well-defined participants with poorly controlled asthma who were taking regular inhaled therapy. The application of detailed induced sputum microbiota characterization from these participants then allowed us to assess the extent to which clinical and inflammatory characteristics independently

### TABLE II. Comparison of patients’ characteristics with α-diversity (Faith’s phylogenetic diversity) and β-diversity (weighted UniFrac distance from centroid) assessed by using the Spearman or Mann-Whitney test

<table>
<thead>
<tr>
<th>Neutrophil percentage</th>
<th>Age</th>
<th>ICS dose</th>
<th>FEV1 (% predicted)</th>
<th>Atopy</th>
<th>Sex</th>
<th>Ever smoked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faith’s diversity</td>
<td>-0.374</td>
<td>-0.309</td>
<td>-0.242</td>
<td>0.193</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P value</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>0.002</td>
<td>0.013</td>
<td>0.32</td>
<td>0.97</td>
</tr>
<tr>
<td>UniFrac distance</td>
<td>0.24</td>
<td>0.015</td>
<td>0.096</td>
<td>0.034</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P value</td>
<td>0.002</td>
<td>0.84</td>
<td>0.22</td>
<td>0.66</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Spearman correlation coefficient (r) and probability values are as indicated.

*\( n = 163 \)

†Assessed by using the Mann-Whitney test.

††\( n = 164 \)

### TABLE III. Multivariate linear regression on α-diversity (Faith’s phylogenetic diversity) and β-diversity (weighted UniFrac distance from centroid) performed on 160 participants

<table>
<thead>
<tr>
<th>B</th>
<th>95% Cl</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faith’s phylogenetic diversity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutrophil percentage</td>
<td>-0.046</td>
<td>-0.07 to -0.02</td>
</tr>
<tr>
<td>Age</td>
<td>-0.036</td>
<td>-0.07 to -0.004</td>
</tr>
<tr>
<td>Sex</td>
<td>0.35</td>
<td>-0.59 to 1.4</td>
</tr>
<tr>
<td>Atopy</td>
<td>-0.77</td>
<td>-1.7 to 0.12</td>
</tr>
<tr>
<td>FEV1 (% predicted)</td>
<td>0.016</td>
<td>-0.01 to 0.04</td>
</tr>
<tr>
<td>ICS dose</td>
<td>-0.001</td>
<td>-0.001 to -0.001</td>
</tr>
<tr>
<td>UniFrac distance from centroid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutrophil percentage</td>
<td>0.14</td>
<td>0.07 to 0.22</td>
</tr>
<tr>
<td>Age</td>
<td>0.056</td>
<td>-0.05 to 0.16</td>
</tr>
<tr>
<td>Sex</td>
<td>-3.9</td>
<td>-7.5 to -3.30</td>
</tr>
<tr>
<td>Atopy</td>
<td>4.8</td>
<td>1.3 to 8.4</td>
</tr>
<tr>
<td>FEV1 (% predicted)</td>
<td>0.049</td>
<td>-0.04 to 0.14</td>
</tr>
<tr>
<td>ICS dose</td>
<td>&lt;0.001</td>
<td>-0.001 to 0.001</td>
</tr>
</tbody>
</table>

*\( n = 163 \)

†Assessed by using the Mann-Whitney test.

††\( n = 164 \)

**Moraxella** taxa in neutrophilic participants, supporting previous findings. This could be interpreted as simply an increased relative abundance of airway pathogens in neutrophilic patients, reflecting neutrophilic influx into the airways during subclinical lower airway infection, with a reciprocal decrease in the relative abundance of commensal taxa. However, a group of common airway taxa correlated negatively with sputum microbiota percentages (Gemella, Porphyromonas, and Streptococcus taxa) and, importantly, even after controlling for overgrowth effects of Haemophilus and Moraxella taxa, significant differences in microbiota composition between neutrophilic and eosinophilic participants were still observed. This finding suggests that separate phenomena contribute to microbial differences between inflammatory phenotypes: the effect of pathogen overgrowth and the selective pressure of airway inflammatory characteristics in the absence of infection. The latter could result in more broad-scale divergence in composition between the neutrophilic and eosinophilic subgroups, in turn contributing to an increased risk of lower airway infection in neutrophilic patients through an increased presence of opportunistic pathogens.

This finding has clear implications for the clinical management of asthma, in which low-dose macrolide and ICS therapies have been shown to influence overgrowth by opportunistic respiratory pathogens and innate immune function, respectively. Furthermore, the relative lack of efficacy of ICSs in patients with noneosinophilic asthma might lead to use of higher doses compared with those used by eosinophilic patients. The combination of underlying differences in airway microbiota (associated with differences in inflammatory phenotype) and inefficacious therapies being used at higher doses might contribute to reduced bacterial diversity, the high concentrations of Proteobacteria seen in the airways of neutrophilic patients, a greater propensity for lung infection, and a further enhancement of the neutrophilic phenotype.

Although strong associations between the neutrophilic phenotype and sputum microbiota composition were found, associations between eosinophil count and microbiota composition were minimal. This contrasts with previous studies reporting increased *Tropheryma* taxon associated with eosinophilia and associations between bronchial biopsy eosinophil count and bacterial composition and reduced bacterial burden associated with type 2–high airway inflammation.

An important strength of our study was its involvement of a large group of well-defined participants with poorly controlled asthma who were taking regular inhaled therapy. The application of detailed induced sputum microbiota characterization from these participants then allowed us to assess the extent to which clinical and inflammatory characteristics independently
associated with variations in airway microbiology by using multivariate linear regression analysis. Multivariate regression identified the sputum neutrophil percentage as the strongest predictor of microbiota variance. However, age, ICS dose, sex, and atopy were also significant independent predictors; conversely, lung function (as measured by FEV₁ percentage) and smoking status were not.

Of particular interest was the finding that increasing age predicted reduced microbiota α-diversity because age has been previously associated with microbiome composition in patients with other chronic respiratory diseases. The airway microbiome of patients with cystic fibrosis is strongly affected by age, which is presumed to relate to the selective effects of increased antibiotic exposure over time and the changing characteristics of the airway environment. It is interesting to speculate that the relationship between neutrophilia and microbiota composition might reflect the effect of age on neutrophilia, suggesting that a tendency toward a neutrophilic phenotype and/or a susceptibility to opportunistic airway infections increases with age in patients with severe asthma.

It is important to recognize a number of limitations of our study. Airway microbiology was assessed based on induced sputum, which, although shown to provide reproducible inflammatory cell levels in patients with moderate-to-severe asthma, only provides an approximation of lower airway microbiology (in common with other lower airway sampling strategies). It is also important to note that induced sputum levels of neutrophils and eosinophils can change frequently and that the relationships between airway microbiology and inflammatory phenotype reported are cross-sectional. Detailed longitudinal analysis is now required to determine how these relationships change with time. 16S rRNA gene amplicon sequence data were subsampled to a level that allowed inclusion of the greatest number of subjects while maintaining sufficient read depth to accurately describe microbiota composition. However, additional analysis with a greater read depth might identify rare taxa that contribute to disease characteristics. Although multivariate regression identified sputum neutrophilia as an independent predictor of microbiota composition, the effects of variation in ICS dose between phenotypes (although nonsignificant) should be noted. Finally, although none of the study participants received antibiotics in the month before recruitment, data on less recent exposure were not available and could have a lasting effect on the lower airway microbiome composition.

The clear relationship between airway inflammatory phenotype and the microbiota highlight the need for studies examining whether asthma treatments should be individualized based on both inflammatory phenotype stratification and lower airway microbiology. There is now a clear need to investigate the extent to which variations in the airway microbiota predict the risk of future asthma exacerbations and to determine whether airway microbiota characterization could be used as a basis for asthma treatment selection.

We thank Heather Powell, Catherine Delahunty, Kellie Fakes, Bridgette Donati, Michelle Gleeson, Erin Harvey, Calida Garside, Gabriele le Brocq, Kelly Steel, Sandra Dowley, Amy Cashmore, Gloria Foxley, Michael Guo, J-Chin Wu, Monique de Pedro, Kirsty Herewane, Miranda Ween, Fungai Madzinga, Melissa McClean, Pamela Fung, Janet Shaw, Joanne McNamara, Kevin Oreto, Peta Grayson, Robyn Jones, Tess Bird, Kerrie Wade, Sara Baum, Chaitali Patel, Michelle Towers, Tina Collins, Melanie Carroll, Alice Chen, Chris Choo, Mary Vukovich, and Kerry Carson for their assistance.

---

**Key messages**

- Lower airway microbiota differs significantly between patients with neutrophilic and eosinophilic asthma characterized based on both increased frequency of pathogenic taxa and divergence in the wider airway microbiota.
- Sputum neutrophilia is the strongest predictor of airway microbiota composition, with age, ICS dose, sex, and atopy also being independent predictors.
- The clear relationships between airway microbiota composition, inflammatory phenotype, and clinical measures suggest microbiota characterization could be a useful contributor to the individualization of asthma treatments.

**REFERENCES**

METHODS

Exclusion criteria

Asthma diagnosis was established by using American Thoracic Society guidelines based on current episodic respiratory symptoms, clinical diagnosis, and evidence of variable airflow obstruction. Participants with asthma were included if stable but symptomatic, despite being prescribed maintenance ICS and long-acting bronchodilator treatment with an Asthma Control Questionnaire score of greater than 7.5. Participants with an FEV₁ of less than 40% of predicted value, current smokers, ex-smokers who had ceased smoking in the previous year, and those with a recent (past 4 weeks) exacerbation or respiratory tract infection were excluded. Those with significant smoking-related air-space disease (ex-smokers with a >10 pack year history and a diffusing capacity/alveolar volume ratio of less than 70% of predicted value or those with a smoking history of greater than 10 pack years and exhaled carbon monoxide >10 ppm) were also excluded. This study was conducted in accordance with the amended Declaration of Helsinki. Local institutional review boards approved the protocol, and written informed consent was obtained from all participants.

Institutional centers

Sputum samples were collected from 8 Australian centers: Hunter Medical Research Institute, Newcastle, Australia; the Prince Charles Hospital, Chermside, Australia; Princess Alexandra Hospital, Woolloongabba, Australia; Royal Adelaide Hospital, Adelaide, Australia; Sir Charles Gardiner Hospital, Nedlands, Australia; the Woolcock Institute of Medical Research, Glebe, Australia; Concord Repatriation General Hospital, Concord, Australia; and Liverpool Hospital, Liverpool, Australia.

Sample collection

All study participants attended a single visit that included assessment of lung function, asthma symptoms, asthma-specific quality of life, medication use, and smoking status. Sputum induction with hypertonic saline (4.5%) was performed, as described previously. Sputum aliquots were stored at −80°C for DNA extraction or dispersed by using dithiothreitol for sputum cell count and assessment of inflammatory subtype determination.

Patient inflammatory phenotyping

Patients` sputum was dispersed by using dithiothreitol, and inflammatory cells were counted as a percentage of total sputum cells. Inflammatory subtype was determined, as described below. Neutrophilic cutoff values were age dependent, as described previously.

Neutrophilic phenotype. Values for the neutrophilic phenotype were as follows: neutrophil percentage (<20 years old), 75.57% or greater; neutrophil percentage (20-40 years old), 61.61% or greater; neutrophil percentage (40-60 years old), 63.25% or greater; and neutrophil percentage (>60 years old), 67.25% or greater.

Eosinophilic phenotype. Values for the eosinophilic phenotype were as follows: eosinophil percentage, 3% or greater.

Paucigranulocytic phenotype. Values for the paucigranulocytic phenotype were as follows: eosinophil percentage, 3% or less; neutrophil percentage (<20 years old), 75.57% or less; neutrophil percentage (20-40 years old), 61.61% or less; neutrophil percentage (40-60 years old), 63.25% or less; and neutrophil percentage (>60 years old), 67.25% and less.

Mixed granulocytic phenotype. Values for the mixed granulocytic phenotype were as follows: eosinophil percentage, 3% or greater; neutrophil percentage (<20 years old), 75.57% or greater; neutrophil percentage (20-40 years old), 61.61% or greater; neutrophil percentage (40-60 years old), 63.25% or greater; and neutrophil percentage (>60 years old), 67.25% or greater.

DNA extraction

DNA extraction was performed on sputum sample aliquots of approximately 100 μL. After the addition of 300 μL of PBS, samples were vortexed for 10 seconds and placed on ice for 2 minutes. Bacterial cells were then pelleted by means of centrifugation at 13,000g for 10 minutes. After removal of supernatant, 300 μL of Tris-EDTA solution (10 mmol/L Tris-HCI and 1 mmol/L EDTA [pH 8.0]; Ambion, Thermo Fisher Scientific, Victoria, Australia), 200 μg of silica/zirconium beads (1:1 of 0.1 μm and 1.0 μm; Bio-Spec Products, Bartlesville, Okla), and a single chrome bead (3.2 mm; Bio-Spec Products) were added to the tube containing the cell pellet. Samples underwent bead beating at 6.5 ms for 60 seconds in a FastPrep-24 Instrument (MP Biomedicals, Burlingame, Calif). The homogenized sample was heated to 90°C for 5 minutes before being cooled on ice for 5 minutes. Lysozyme (ROCHE, Thermo Fisher Scientific, Victoria, Australia) and lysostaphin (Sigma-Aldrich, St Louis, Mo) were then added to a final concentration of 2 and 0.1 mg/mL, respectively, and samples were incubated at 37°C for 2 hours. Proteinase K (Fermentas, Thermo Fisher Scientific, Victoria, Australia) and sodium dodecyl sulfate (Sigma-Aldrich) were then added to a final concentration of 1.2 mg/mL and 1.5% (wt/vol), respectively. After incubation at 30 minutes at 56°C, 40 μL of 5 mol/L sodium chloride and 450 μL of phenol/chloroform/isoamyl alcohol (25:24:1; saline buffered at pH 8.0; Sigma-Aldrich) were added, and samples were vortexed for 30 seconds. The aqueous organic layers were separated by means of centrifugation at 13,000g for 10 minutes, and 400 μL of the aqueous layer was transferred to a new microfuge tube. DNA was recovered by using an EZ-10 Spin column in accordance with the manufacturer’s instructions (Bio Basic, Markham, Ontario, Canada) after precipitation by addition of 10 mol/L ammonium acetate and 99% ethanol (Sigma Aldrich) in a 1:10 and 1:1 ratio with sample volume, respectively. DNA was eluted in 50 μL of UltraPure DNase/RNase-free distilled water (Gilbo, Thermo Fisher Scientific, Victoria, Australia), and stored at −80°C before analysis.

16S rRNA gene amplicon sequencing

The V1-3 hypervariable region of the bacterial 16S rRNA gene was amplified from sputum DNA by using modified primers 27F (5'-TCGTCCGGCAACGCTACAGATGTAGTTATCAAGACGACAGGGTGTGTACMTGCTTCAG-3') and 519R (5'-GTCTCGTGGGCTCAGGGAATGATTTAAAGAGCAGGTNTTACNGCGGCKCTG-3'), with Illumina adapter overhang sequences as indicated by underlining. Amplicons were generated, cleaned, indexed, and sequenced according to the Illumina MiSeq 16S Metagenomic Sequencing Library Preparation Protocol (http://support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html) with certain modifications. Briefly, an initial PCR reaction contained at least 12.5 ng of DNA, 5 μL of forward primer (1 μmol/L), 1.5 μL of reverse primer (1 μmol/L), and 12.5 μL of 2× KAPA HiFi Hotstart ReadyMix (KAPA Biosystems, Wilmington, Mass) in a total volume of 25 μL. The PCR reaction was performed on a Veriti 96-well Thermal Cycler (Life Technologies, Grand Island, NY) by using the following program: 95°C for 3 minutes, followed by 25 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 30 seconds and a final extension step at 72°C for 5 minutes. Samples were multiplexed by using a dual-index approach with the Nextera XT Index kit (Illumina), according to the manufacturer’s instructions. The final library was paired-end sequenced at 2 × 300 bp by using a MiSeq Reagent Kit v3 on the Illumina MiSeq platform. Sequencing was performed at the David R Gunn Genomics Facility, South Australian Health and Medical Research Institute.

16S rRNA gene qPCR

The approximate 16S rRNA gene copy number was assessed by using qPCR with the 16S rRNA universal primers B331F (5'-TCTTACGCGGACGCGAT-3') and B797R (5'-GGACTACAGGGTGATCCTAATCTGTTT-3') and Platinum SYBR Green (Thermo Fisher Scientific), as previously described.13 Reactions were performed in duplicates, and averages were taken. Sample total bacterial copy number was calculated per microliter of DNA eluate against a standard curve of a known bacterial copy number.

Sequence data processing

The Quantitative Insights Into Microbial Ecology (version 1.8.0) software was used to analyze the 16S rRNA sequence generated from
paired-end amplicon sequencing by using a bioinformatics pipeline, as previously described. Briefly, barcoded forward and reverse sequencing reads were quality filtered and merged with Paired-End reAd mergeR (version 0.9.6). Chimeras were detected and filtered from the paired-end reads by using USEARCH (version 6.1) against the 97% clustered representative sequences from the Greengenes database (versus 13.8). OTUs were assigned to the reads by using an open reference approach with the UCLUST algorithm (version 1.2.29q) against the SILVA database (release 111, July 2012), which was clustered at 97% identity. Spurious OTUs were then removed systematically by using previous reports of common laboratory sequencing contaminants. A minimum subsampling depth of 1732 reads was then selected for all samples.

Where taxa assignment did not classify to the family or genus level, OTU reference sequences (accounting for >99% of OTU reads) were separately aligned by using the SILVA Incremental Aligner (https://www.arb-silva.de/), which uses SILVA, RDF, Greengenes, LTP, and EMBL sequence collections. If the alignments identified taxa to the genus level and at greater than 99% similarity, they replaced the previous taxon assignment. This occurred for *Streptococcus II*, which was previously incorrectly assigned as Clostridiales; Other; *Other. Streptococcus I* refers to the OTU cluster, which was assigned as *Streptococcus* taxon during initial assignment.

**Diversity measurements and statistical analyses**

The Bray-Curtis matrix was calculated based on sample-normalized, square root–transformed relative taxon abundance. Principal coordinate analysis was used to visualize clustering of samples based on their similarity matrices, with PCO1 and PCO2 coordinates and group centroids plotted by using the ggplot2 package of R statistical software. Distance from centroid was calculated, as previously described, by using PRIMER. PERMANOVA on the β-diversity matrices was used to test the null hypothesis of no difference among a priori–defined groups with the PERMANOVA + add-on package for PRIMER. The test was computed by using unrestricted permutation of raw data with 9999 random permutations at a significance level of .01.

**Taxon dispersion**

Variation in microbiota composition at the genus level was assessed by using multiple approaches. First, taxa that contributed to the overall variation between asthma phenotypes were identified by using SIMPER analysis in PRIMER. Subsequently, the abundance of the 13 highest ranked taxa (accounting for 50% of the dissimilarity between neutrophilic and eosinophilic groups) were used to generate a heat map with the ggplot2 package of R statistical software. Hierarchical clustering of the taxa was performed on Bray-Curtis dissimilarity and clustered by using the single linkage method. Dominance of *Haemophilus* and *Moraxella* taxa was determined when the relative abundance of each taxa exceeded 40%. This cutoff was selected based on the distribution of the relative abundance, where a clear distinction between samples with greater than 40% and less than 40% was evident, suggesting overgrowth of these taxa.

Second, strong taxon-taxon correlations were identified by using SparCC where absolute taxon abundances were bootstrapped 100 times to generate correlation *P* values. Networks were then generated from selected *r* > 0.25 or *r* < -0.25, and *P* values (*P* ≤ 0.1) were generated by using Cytoscape (version 3.4.0).

Two approaches were used to investigate the effect of pathogen overgrowth on microbiota composition. First, in samples in which *Haemophilus* or *Moraxella* taxa were the dominant taxa and represented 40% or greater of total reads, their relative abundance was adjusted to the mean value for the study cohort, and the remaining relative abundance measures were rescaled, as described previously. PERMANOVA analyses were then performed on the rescaled data.

Second, pairwise comparisons between neutrophilic and eosinophilic samples were performed by using the phyloseq R package with the DEseq2 extension based on count data. *P* values were corrected by using the Benjamini-Hochberg false discovery rate procedure, and a corrected α value cutoff of less than 0.05 was used for inclusion.

**REFERENCES**

E1. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. Am Rev Respir Dis 1987;136:225-44.