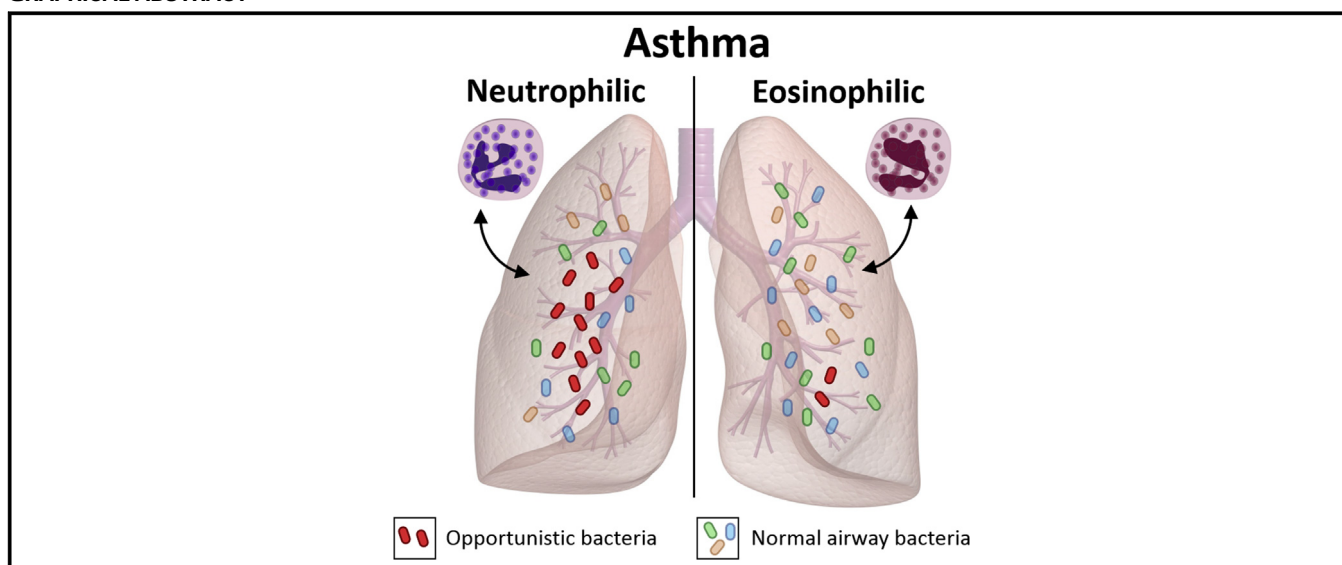# Inflammatory phenotypes in patients with severe asthma are associated with distinct airway microbiology

CrossMark

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,e] Paul N. Reynolds, MBBS, PhD,[f,g] Sandra Hodge, PhD,[f,g]
Alan L. James, FRACP, PhD,[h,i] 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]‡   *Adelaide, St Lucia, Chermside, Wooloongabba, Nedlands,
Crawley, Newtown, North Ryde, Concord, Liverpool, Glebe, Callaghan, and Sydney, Australia*

GRAPHICAL ABSTRACT

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 α- and β-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 (α-diversity: Spearman $r = -0.374$, $P < .001$; β-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.)

Evidence-based patient stratification is required to address the heterogeneity of asthma severity and treatment responsiveness. Asthma phenotype based on characteristics of airway inflammation are increasingly recognized as an important prognostic indicator.[1] In addition to an allergen-induced, T_H2 lymphocyte, IL-5–mediated, eosinophilic inflammatory response, asthma can also occur in the absence of eosinophilic inflammation (termed noneosinophilic asthma).[2,3] 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.[2]

Unlike the relatively well-defined mechanisms that result in eosinophilic airway inflammation, those leading to noneosinophilic asthma, particularly neutrophilic asthma, remain relatively poorly understood.[4] Furthermore, although noneosinophilic phenotypes occur across the spectrum of asthma severity,[1,2,5] they typically respond poorly to corticosteroids.[2,5] Inflammatory phenotypes have also been shown to differ with respect to airway

Abbreviations used	
ICS:	Inhaled corticosteroid
OTU:	Operational taxonomic unit
PERMANOVA:	Permutational multiple ANOVA
qPCR:	Quantitative PCR
SIMPER:	Similarity of percentages

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[6] or culture-independent[5] approaches and have reduced airway bacterial diversity.[7] Given that airway microbiota composition is associated with the degree of airway hyperresponsiveness among patients with suboptimally controlled asthma,[8] 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 physico-chemical 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.[9] 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.[10] 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.[5,11-14] 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.[3,15] 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.[16] Spurious operational taxonomic units (OTUs) were removed systematically by using previous reports of common laboratory sequencing contaminants.[17] Bacterial burden was quantified with quantitative PCR (qPCR) for the 16S rRNA gene.[18] 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 equitability of detected taxa). β-Diversity (intersample variance) was determined by using 2 approaches: weighted UniFrac similarity (which accounts for phylogenetic distance) and square root–transformed Bray-Curtis similarity (based on the relative abundance of taxa alone). Both α- and β-diversity measures were calculated with either QIIME (version 1.8.0) or PRIMER (version 6; PRIMER-E, Plymouth, United Kingdom) software.

Continuous data were tested for nonnormality, including skewness and kurtosis, by using the D'Agostino-Pearson omnibus test. Kruskal-Wallis 1-way ANOVA with the Dunn *post hoc* test was used for multiple comparisons of nonnormally distributed data, the Mann-Whitney *U* test was used for pairwise comparisons, the χ^2 test was used for categorical data, and the Spearman test was used for correlations (GraphPad Prism, version 7.01; GraphPad Software, La Jolla, Calif).

Multivariate linear regression was performed with Faith's phylogenetic diversity and UniFrac distance from centroid as 2 dependent variables reflecting aspects of diversity (SPSS software, version 23.0; IBM, Armonk, NY). Covariates were selected *a priori* and included in the model based on a significant correlation with either dependent variable. CIs were obtained by means of bootstrapping, resampling 1000 times. Covariates were tested for collinearity by using variance inflation factors.

Taxon dispersion

Variation in microbiota composition between groups was assessed by using similarity of percentages (SIMPER) analysis in PRIMER and SparCC, from which correlations ($r \geq 0.25$ or $r \leq -0.25$) and P values ($P \leq .01$) were visualized by means of network analysis with Cytoscape (version 3.4.0). Pathogen overgrowth adjustment is described in the Methods section in this article's Online Repository.

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).[3,15] 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 6 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 FEV_1 percent predicted ($P = .035$) and FEV_1/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 ($r = -0.374$, $P < .0001$; Fig 1, *B*) but not with the sputum eosinophil percentage ($r = 0.146$, $P = .060$; 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

	Neutrophilic	Eosinophilic	Paucigranulocytic	Mixed granulocytic	*P* value
No.	14	84	60	9	
Age (y), mean (SD)	59.8 (13.9)	57.2 (15.2)	55.8 (13.8)	56.8 (17.8)	.817
Male sex, no. (%)	9 (64.3)	30 (35.7)	23 (38.3)	6 (66.7)	.084
Atopic, no. (%)	11 (78.6)	62 (76.5), n = 81	52 (86.7)	6 (66.7)	.305
Previous smoker, no. (%)	6 (42.9)	32 (38.1)	17 (28.3)	4 (44.4)	.491
Smoking pack years, median (Q1, Q3)	22.0 (18.6, 27.5)	5.3 (1.2, 15.9)	5.0 (1.3, 30.0)	4.4 (0.9, 63.6)	.379
Duration of asthma (y), median (Q1, Q3)	33.8 (3.5, 48.3)	36.3 (19.7, 49.2)	33.7 (14.2, 54.4)	53.6 (32.6, 60.4)	.205
FEV_1 (% predicted), mean (SD)	70.3 (18.2)	70.0 (17.9)	78.7 (19.2)*	68.4 (17.3)	.035
FVC (% predicted), mean (SD)	81.4 (11.6)	83.6 (16.4)	85.1 (15.0)	82.9 (11.7)	.854
FEV_1/FVC (%), mean (SD)	65.8 (14.5)	64.9 (12.2)	71.3 (12.2)*	63.9 (12.7)	.013
ACQ6 score, mean (SD)	2.1 (1.2)	1.9 (0.9)	1.6 (0.8)	1.5 (0.6)	.210
GINA treatment step					.058
1	1 (7.1)	2 (2.5)	0 (0)	0 (0)	
2	0 (0)	0 (0)	0 (0)	0 (0)	
3	0 (0)	10 (12.4)	12 (20.3)	2 (22.2)	
4	11 (78.6)	67 (82.7)	47 (80.0)	7 (77.8)	
5	2 (14.3)	2 (2.5)	0 (0)	0 (0)	
ICS dose (μg), median (Q1, Q3)	2000 (1280, 2000), n = 13	1000 (800, 2000), n = 82	1000 (800, 2000), n = 59	1600 (1000, 2000)	.254
Total cell count (× 10^6/mL [Q1, Q3])	9.8 (7.6, 12.7)*†	3.6 (1.9, 7.6)	3.74 (2.0, 7.6)	8.64 (5.22, 11.34)	<.001
Viability (%)	89.2 (73.5, 93.0)*†	69.8 (52.2, 80.7)	68.9 (55.7, 79.5)	90.0 (84.4, 93.7)*†	<.001
Neutrophils (%)	75.0 (68.80, 84.00)*†	27.13 (14.38, 41.00)	34.13 (12.63, 49.25)	77.00 (71.25, 77.50)*†	<.001
Eosinophils (%)	0.63 (0.50, 1.00)*	6.88 (3.63, 18.13)	0.25 (0.00, 1.00)*	6.75 (3.25, 12.25)†‡	<.001
Macrophages (%)	20.75 (14.50, 30.25)*†	52.75 (38.25, 71.00)	53.34 (43.50, 75.63)	16.00 (15.50, 20.25)*†	<.001
Lymphocytes (%)	0.63 (0.25, 1.00)	1.13 (0.25, 2.25)	0.88 (0.25, 2.63)	0.25 (0.00, 0.38)*†	.018
Columnar epithelial cells (%)	0.25 (0.00, 1.25)*†	2.17 (0.75, 6.00)	4.13 (1.21, 8.50)	0.50 (0.25, 1.50)†	<.001
Squamous cells (%)	1.11 (0.50, 3.85)*†	5.99 (2.32, 16.23)	6.18 (2.92, 11.31)	3.38 (0.25, 8.05)	.002

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 ($\chi^2 = 25.5$, $P < .0001$; 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*
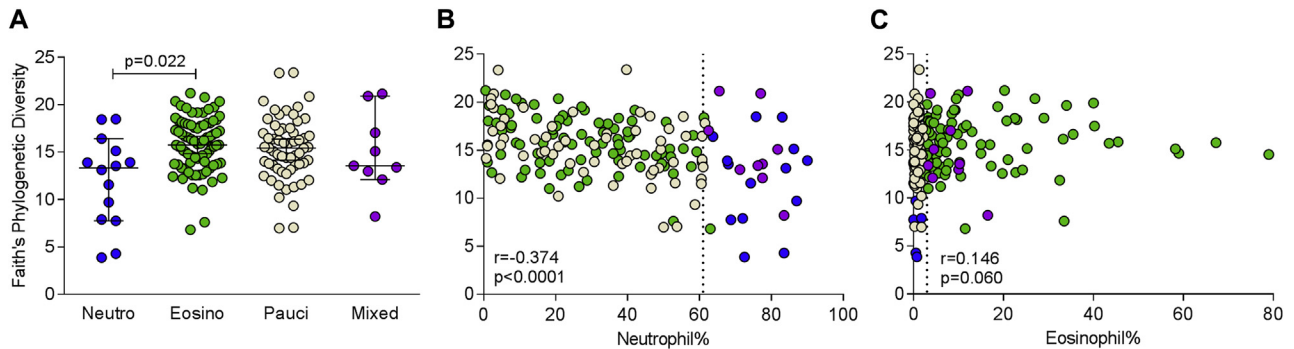
FIG 1. Faith's phylogenetic diversity is significantly associated with sputum neutrophilia but not eosinophilia. **A,** Patients grouped by asthma phenotype. **B,** Neutrophil percentage. The *dotted line* at 61% neutrophils indicates the phenotype cutoff point. **C,** Eosinophil percentage. The *dotted 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*).

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 FEV_1 percentage (Table II). Conversely, weighted UniFrac distance from centroid significantly positively correlated with sputum neutrophil percentage and was significantly different
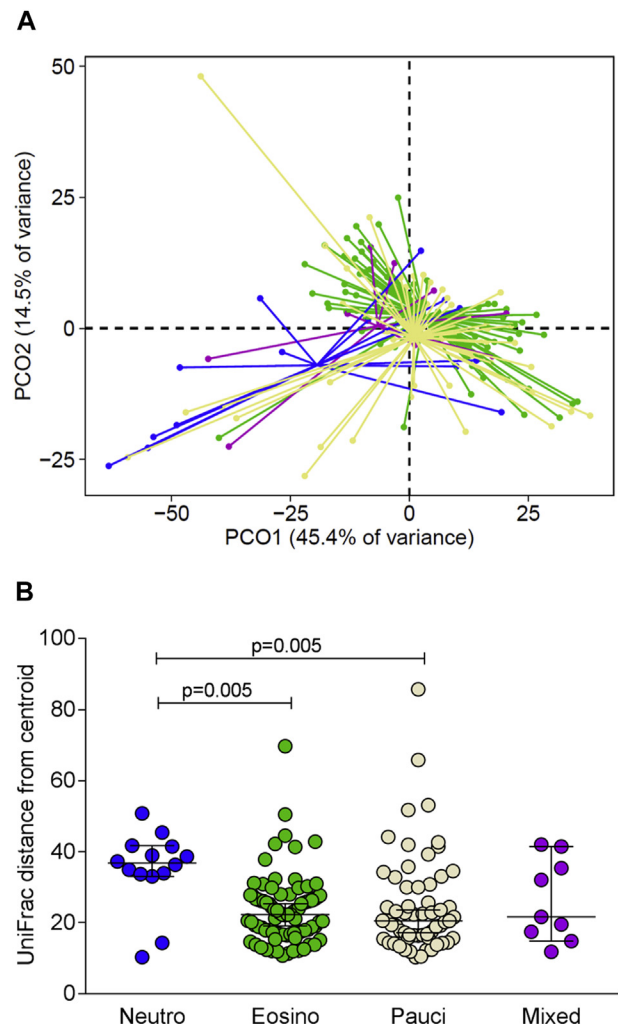

FIG 2. Microbiota dispersion grouped by asthma phenotype. **A,** Principal coordinate analysis. The first 2 principal coordinates are plotted on the *x*- and *y*-axes, respectively (representing 59.9% of the total variation). **B,** Distance from centroid. Statistical significance was assessed by using Kruskal-Wallis 1-way ANOVA with the Dunn *post hoc* test.
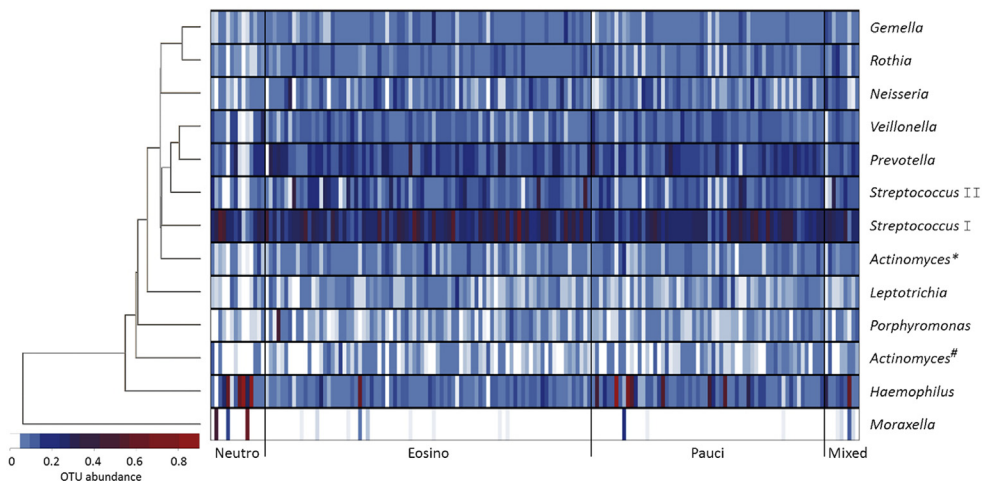
FIG 3. Relative abundance of discriminant taxa among asthma phenotypes. The 13 taxa that collectively contribute to approximately 50% of variance among phenotypes, as determined by using SIMPER analysis. Clustering shows the similarity relationship of genera based on Bray-Curtis similarity distance and the single linkage hierarchical clustering method. *Taxa assigned *Actinomyces* species uncultured bacteria. #Taxa assigned *Actinomyces* species oral clone DR002.
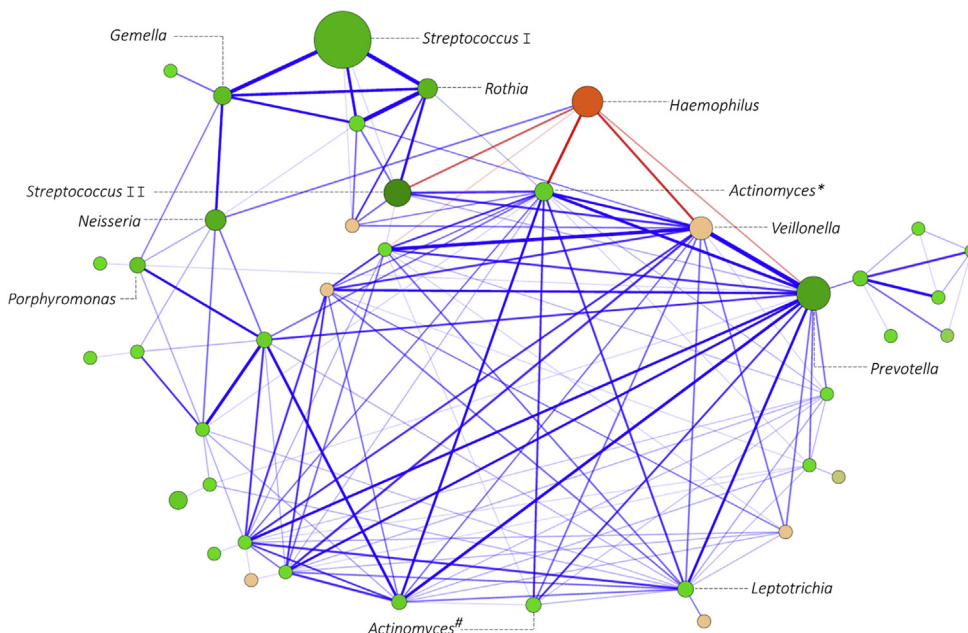

FIG 4. Bacterial network analysis of the asthma cohort. Each edge represents a significant correlation colored to indicate either positivity (blue) or negativity (red). Edge width and transparency are proportional to the absolute value of the correlation coefficient. Node size is proportional to mean 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.

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 ($P = .030$ [95% CI = −0.07 to −0.004] and $P = .042$ [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
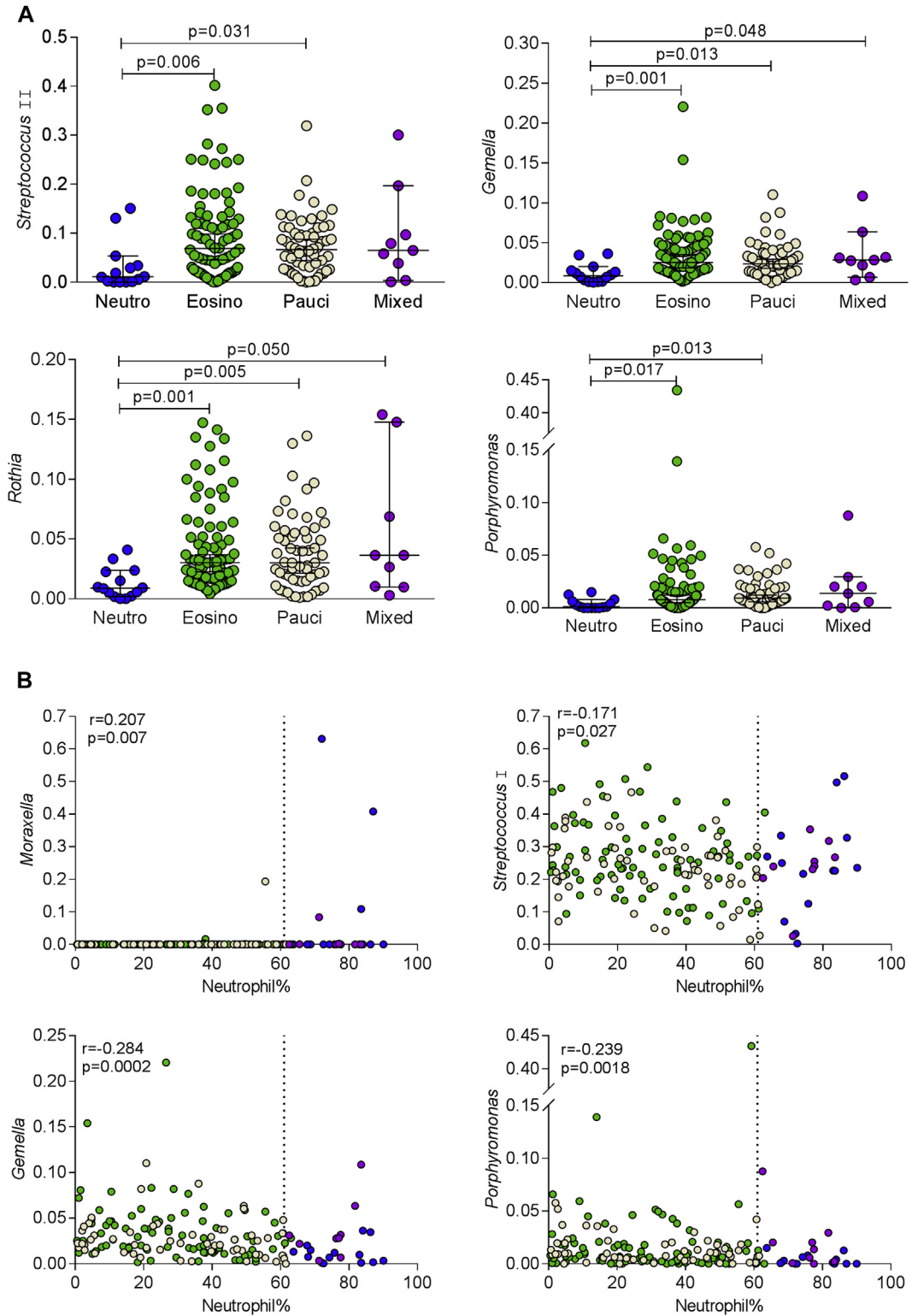
FIG 5. Taxa distribution differs by sputum neutrophilia. **A,** Taxa that significantly differ by patient inflammatory phenotype. **B,** Significant correlations between taxa and neutrophil percentages. Colors represent asthma phenotype based on neutrophilia or eosinophilia: blue is greater than 61% 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*).

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

		Neutrophil percentage	Age	ICS dose*	FEV₁ (% predicted)	Atopy†‡	Sex†	Ever smoked†
Faith's diversity	r	−0.374	−0.309	−0.242	0.193			
	P value	<0.001	<0.001	0.002	0.013	0.32	0.97	0.53
UniFrac distance	r	0.24	0.015	0.096	0.034			
	P value	0.002	0.84	0.22	0.66	0.019	0.003	0.70

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

	B	95% CI	P value
Faith's phylogenetic diversity			
Neutrophil percentage	−0.046	−0.07 to −0.02	.002
Age	−0.036	−0.07 to −0.004	.030
Sex	0.35	−0.59 to 1.4	.49
Atopy	−0.77	−1.7 to 0.12	.10
FEV₁ (% predicted)	0.016	−0.01 to 0.04	.23
ICS dose	−0.001	−0.001 to −0.001	.042
UniFrac distance from centroid			
Neutrophil percentage	0.14	0.07 to 0.22	<.001
Age	0.056	−0.05 to 0.16	.30
Sex	−3.9	−7.5 to −0.30	.039
Atopy	4.8	1.3 to 8.4	.018
FEV₁ (% predicted)	0.049	−0.04 to 0.14	.27
ICS dose	<0.001	−0.001 to 0.001	.56

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 infections in asthmatic patients[5,7] and in those with other respiratory disorders,[19-21] as well as with the effects of exposure to antibiotics.[22,23] Importantly, none of the study participants reported clinical features of respiratory tract infection or had antibiotic therapy during the preceding month.

We further assessed α-diversity metrics relative to continuous neutrophil and eosinophil count data as an alternative to categorical inflammatory phenotypes. Significant correlations were observed between sputum neutrophil percentages and each assessed α-diversity metric, with no significant interactions between any diversity metric and sputum eosinophil percentages, strongly suggesting that decreased microbiota richness, evenness, and diversity are associated with airway neutrophilia. Analysis of sputum microbiota β-diversity (intersample similarity) also demonstrated substantial differences between patients with neutrophilic airway inflammation and those with other inflammatory phenotypes, which is consistent with the stochastic overgrowth of complex commensal communities by individual opportunistic pathogens.[24]

We identified bacterial taxa that contributed to observed differences in microbiota composition between inflammatory phenotypes, namely high abundance of *Haemophilus* and *Moraxella* taxa in neutrophilic participants, supporting previous findings.[5] 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 neutrophil 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 2 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.[6,14,25] Furthermore, the relative lack of efficacy of ICSs in patients with noneosinophilic asthma[2] 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,[14] the high concentrations of Proteobacteria seen in the airways of neutrophilic patients,[7] a greater propensity for lung infection, and a further enhancement of the neutrophilic phenotype.[6]

Although strong associations between the neutrophilic phenotype and sputum microbiota composition were found, associations between eosinophil counts and microbiota composition were minimal. This contrasts with previous studies reporting increased *Tropheryma* taxon associated with eosinophilia[7] and associations between bronchial biopsy eosinophil count and bacterial composition[11] and reduced bacterial burden associated with type 2–high airway inflammation.[26]

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 FEV_1 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,[27] which is presumed to relate to the selective effects of increased antibiotic exposure over time and the changing characteristics of the airway environment.[9] It is interesting to speculate that the relationship between neutrophilia and microbiota composition might reflect the effect of age on neutrophilia,[15] 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,[28,29] only provides an approximation of lower airway microbiology (in common with other lower airway sampling strategies).[30] It is also important to note that induced sputum levels of neutrophils and eosinophils can change frequently[31,32] 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.[22]

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.

Key messages

- **Lower airway microbiology 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

1. Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. Nat Med 2012;18:716-25.
2. Pavord ID, Brightling CE, Woltmann G, Wardlaw AJ. Non-eosinophilic corticosteroid unresponsive asthma. Lancet 1999;353:2213-4.
3. Simpson JL, Scott R, Boyle MJ, Gibson PG. Inflammatory subtypes in asthma: assessment and identification using induced sputum. Respirology 2006;11:54-61.
4. Pelaia G, Vatrella A, Busceti MT, Gallelli L, Calabrese C, Terracciano R, et al. Cellular mechanisms underlying eosinophilic and neutrophilic airway inflammation in asthma. Mediat Inflamm 2015;2015:879783.
5. Green BJ, Wiriyachaiporn S, Grainge C, Rogers GB, Kehagia V, Lau L, et al. Potentially pathogenic airway bacteria and neutrophilic inflammation in treatment resistant severe asthma. PLoS One 2014;9:e100645.
6. Essilfie AT, Simpson JL, Dunkley ML, Morgan LC, Oliver BG, Gibson PG, et al. Combined Haemophilus influenzae respiratory infection and allergic airways disease drives chronic infection and features of neutrophilic asthma. Thorax 2012;67:588-99.
7. Simpson JL, Daly J, Baines KJ, Yang IA, Upham JW, Reynolds PN, et al. Airway dysbiosis: *Haemophilus influenzae* and *Tropheryma* in poorly controlled asthma. Eur Respir J 2016;47:792-800.
8. Huang YJ, Nelson CE, Brodie EL, Desantis TZ, Baek MS, Liu J, et al. Airway microbiota and bronchial hyperresponsiveness in patients with suboptimally controlled asthma. J Allergy Clin Immunol 2011;127:372-81, e1-3.
9. Rogers GB, Hoffman LR, Carroll MP, Bruce KD. Interpreting infective microbiota: the importance of an ecological perspective. Trends Microbiol 2013;21:271-6.
10. McKeever T, Harrison TW, Hubbard R, Shaw D. Inhaled corticosteroids and the risk of pneumonia in people with asthma: a case-control study. Chest 2013;144:1788-94.
11. Huang YJ, Nariya S, Harris JM, Lynch SV, Choy DF, Arron JR, et al. The airway microbiome in patients with severe asthma: associations with disease features and severity. J Allergy Clin Immunol 2015;136:874-84.
12. Simpson JL, Carroll M, Yang IA, Reynolds PN, Hodge S, James AL, et al. Reduced antiviral interferon production in poorly controlled asthma is associated with neutrophilic inflammation and high-dose inhaled corticosteroids. Chest 2016;149:704-13.
13. Zhang Q, Cox M, Liang Z, Brinkmann F, Cardenas PA, Duff R, et al. Airway microbiota in severe asthma and relationship to asthma severity and phenotypes. PLoS One 2016;11:e0152724.
14. Denner DR, Sangwan N, Becker JB, Hogarth DK, Oldham J, Castillo J, et al. Corticosteroid therapy and airflow obstruction influence the bronchial microbiome, which is distinct from that of bronchoalveolar lavage in asthmatic airways. J Allergy Clin Immunol 2016;137:1398-405.e3.
15. Brooks CR, Gibson PG, Douwes J, Dalen CJV, Simpson JL. Relationship between airway neutrophilia and ageing in asthmatics and non-asthmatics. Respirology 2013;18:857-65.
16. Jervis-Bardy J, Leong LE, Marri S, Smith RJ, Choo JM, Smith-Vaughan HC, et al. Deriving accurate microbiota profiles from human samples with low bacterial content through post-sequencing processing of Illumina MiSeq data. Microbiome 2015;3:19.

17. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol 2014;12:1-12.
18. Nadkarni MA, Martin FE, Jacques NA, Hunter N. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. Microbiology 2002;148:257-66.
19. Cuthbertson L, Rogers GB, Walker AW, Oliver A, Green LE, Daniels TW, et al. Respiratory microbiota resistance and resilience to pulmonary exacerbation and subsequent antimicrobial intervention. ISME J 2016;10:1081-91.
20. Hofstra JJ, Matamoros S, van de Pol MA, de Wever B, Tanck MW, Wendt-Knol H, et al. Changes in microbiota during experimental human rhinovirus infection. BMC Infect Dis 2015;15:336.
21. Pettigrew MM, Gent JF, Kong Y, Wade M, Gansebom S, Bramley AM, et al. Association of sputum microbiota profiles with severity of community-acquired pneumonia in children. BMC Infect Dis 2016;16:317.
22. Daniels TW, Rogers GB, Stressmann FA, van der Gast CJ, Bruce KD, Jones GR, et al. Impact of antibiotic treatment for pulmonary exacerbations on bacterial diversity in cystic fibrosis. J Cystic Fibros 2013;12:22-8.
23. Huang YJ, Sethi S, Murphy T, Nariya S, Boushey HA, Lynch SV. Airway microbiome dynamics in exacerbations of chronic obstructive pulmonary disease. J Clin Microbiol 2014;52:2813-23.
24. Tarabichi Y, Li K, Hu S, Nguyen C, Wang X, Elashoff D, et al. The administration of intranasal live attenuated influenza vaccine induces changes in the nasal microbiota and nasal epithelium gene expression profiles. Microbiome 2015;3:74.
25. Rogers GB, Bruce KD, Martin ML, Burr LD, Serisier DJ. The effect of long-term macrolide treatment on respiratory microbiota composition in non-cystic fibrosis bronchiectasis: an analysis from the randomised, double-blind, placebo-controlled BLESS trial. Lancet Respir Med 2014;2:988-96.
26. Durack J, Lynch SV, Nariya S, Bhakta NR, Beigelman A, Castro M, et al. Features of the bronchial bacterial microbiome associated with atopy, asthma, and responsiveness to inhaled corticosteroid treatment. J Allergy Clin Immunol 2017;140:63-75.
27. Zhao J, Schloss PD, Kalikin LM, Carmody LA, Foster BK, Petrosino JF, et al. Decade-long bacterial community dynamics in cystic fibrosis airways. Proc Natl Acad Sci U S A 2012;109:5809-14.
28. Bacci E, Cianchetti S, Carnevali S, Bartoli ML, Dente FL, Di Franco A, et al. Induced sputum is a reproducible method to assess airway inflammation in asthma. Mediat Inflamm 2002;11:293-8.
29. Rossall MR, Cadden PA, Molphy SD, Plumb J, Singh D. Repeatability of induced sputum measurements in moderate to severe asthma. Respir Med 2014;108:1566-8.
30. Rogers GB, Skelton S, Serisier DJ, van der Gast CJ, Bruce KD. Determining cystic fibrosis-affected lung microbiology: comparison of spontaneous and serially induced sputum samples by use of terminal restriction fragment length polymorphism profiling. J Clin Microbiol 2010;48:78-86.
31. Hancox RJ, Cowan DC, Aldridge RE, Cowan JO, Palmay R, Williamson A, et al. Asthma phenotypes: Consistency of classification using induced sputum. Respirology 2012;17:461-6.
32. Suarez-Cuartin G, Crespo A, Mateus E, Torrejon M, Giner J, Belda A, et al. Variability in asthma inflammatory phenotype in induced sputum. Frequency and causes. Arch Bronconeumol 2016;52:76-81.

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.[E1] Participants with asthma were included if stable but symptomatic, despite being prescribed maintenance ICS and long-acting bronchodilator treatment with an Asthma Control Questionnaire 6 score of greater than 0.75.[E2]

Participants with an FEV_1 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 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 Gairdner 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,[E3] medication use, and smoking status. Sputum induction with hypertonic saline (4.5%) was performed, as described previously.[E4] Sputum aliquots were stored at −80°C for DNA extraction or dispersed by using dithiothreitol for sputum cell count assessment and inflammatory subtype determination.[E5]

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.[E6,E7]

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-HCl 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 m/s 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 1 hour. 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 (Gibco, 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′-<u>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG</u>AGRGTTTGATCMTGGCTCAG-3′) and 519R (5′-<u>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG</u>GTNTTACNCGGCKGCTG-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), 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, Grans 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′-TCCTACGGGAGGCAGCAGT-3′) and B797R (5′-GGACTACCAGGGTATCTAATCCTGTT-3′) and Platinum SYBR Green (Thermo Fisher Scientific), as previously described.[E8] 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)[E9] software was used to analyze the 16S rRNA sequence generated from

paired-end amplicon sequencing by using a bioinformatics pipeline, as previously described.[E10] Briefly, barcoded forward and reverse sequencing reads were quality filtered and merged with Paired-End reAd mergeR (version 0.9.6).[E11] Chimeras were detected and filtered from the paired-end reads by using USEARCH (version 6.1)[E12] against the 97% clustered representative sequences from the Greengenes database (versus 13.8).[E13] OTUs were assigned to the reads by using an open reference approach with the UCLUST algorithm (version 1.2.22q) against the SILVA database (release 111, July 2012),[E14] which was clustered at 97% identity. Spurious OTUs were then removed systematically by using previous reports of common laboratory sequencing contaminants.[E15] 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, RDP, 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.[E16] Distance from centroid was calculated, as previously described, by using PRIMER.[E17] PERMANOVA[E18] 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.[E16] 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 to 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,[E19] where absolute taxon abundances were bootstrapped 100 times to generate correlation *P* values. Networks were then generated from selected correlations ($r \geq 0.25$ or $r \leq -0.25$), and *P* values ($P \leq .01$) 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.[E20] 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[E21] with the DEseq2[E22] 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.

E2. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. Eur Respir J 1999;14:902-7.

E3. Juniper EF, Guyatt GH, Ferrie PJ, Griffith LE. Measuring quality of life in asthma. Am Rev Respir Dis 1993;147:832-8.

E4. Gibson PG, Wlodarczyk JW, Hensley MJ, Gleeson M, Henry RL, Cripps AW, et al. Epidemiological association of airway inflammation with asthma symptoms and airway hyperresponsiveness in childhood. Am J Respir Crit Care Med 1998; 158:36-41.

E5. Simpson JL, Phipps S, Baines KJ, Oreo KM, Gunawardhana L, Gibson PG. Elevated expression of the NLRP3 inflammasome in neutrophilic asthma. Eur Respir J 2014;43:1067-76.

E6. Brooks CR, Gibson PG, Douwes J, Dalen CJV, Simpson JL. Relationship between airway neutrophilia and ageing in asthmatics and non-asthmatics. Respirology 2013;18:857-65.

E7. Simpson JL, Scott R, Boyle MJ, Gibson PG. Inflammatory subtypes in asthma: assessment and identification using induced sputum. Respirology 2006;11:54-61.

E8. Nadkarni MA, Martin FE, Jacques NA, Hunter N. Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. Microbiology 2002;148:257-66.

E9. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010;7:335-6.

E10. Jervis-Bardy J, Leong LE, Marri S, Smith RJ, Choo JM, Smith-Vaughan HC, et al. Deriving accurate microbiota profiles from human samples with low bacterial content through post-sequencing processing of Illumina MiSeq data. Microbiome 2015;3:19.

E11. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics 2014;30:614-20.

E12. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 2010;26:2460-1.

E13. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J 2012;6:610-8.

E14. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 2013;41(Database issue):D590-6.

E15. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol 2014;12:1-12.

E16. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2009. http://ggplot2.org.

E17. Anderson MJ, Ellingsen KE, McArdle BH. Multivariate dispersion as a measure of beta diversity. Ecol Lett 2006;9:683-93.

E18. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol 2001;26:2-46.

E19. Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. PLoS Comp Biol 2012;8:e1002687.

E20. Rogers GB, van der Gast CJ, Serisier DJ. Predominant pathogen competition and core microbiota divergence in chronic airway infection. ISME J 2015;9:217-25.

E21. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 2013;8:e61217.

E22. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.
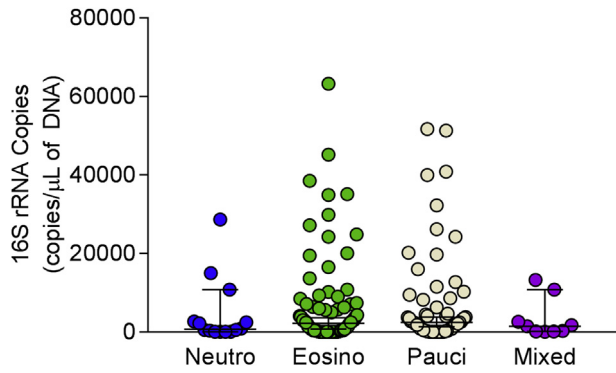
FIG E1. Bacterial burden, as assessed based on 16S rRNA gene copy number. *Bars* show medians ± 95% CIs. Statistical significance was assessed by using Kruskal-Wallis 1-way ANOVA with the Dunn *post hoc* test. No significant difference between phenotypes was seen.
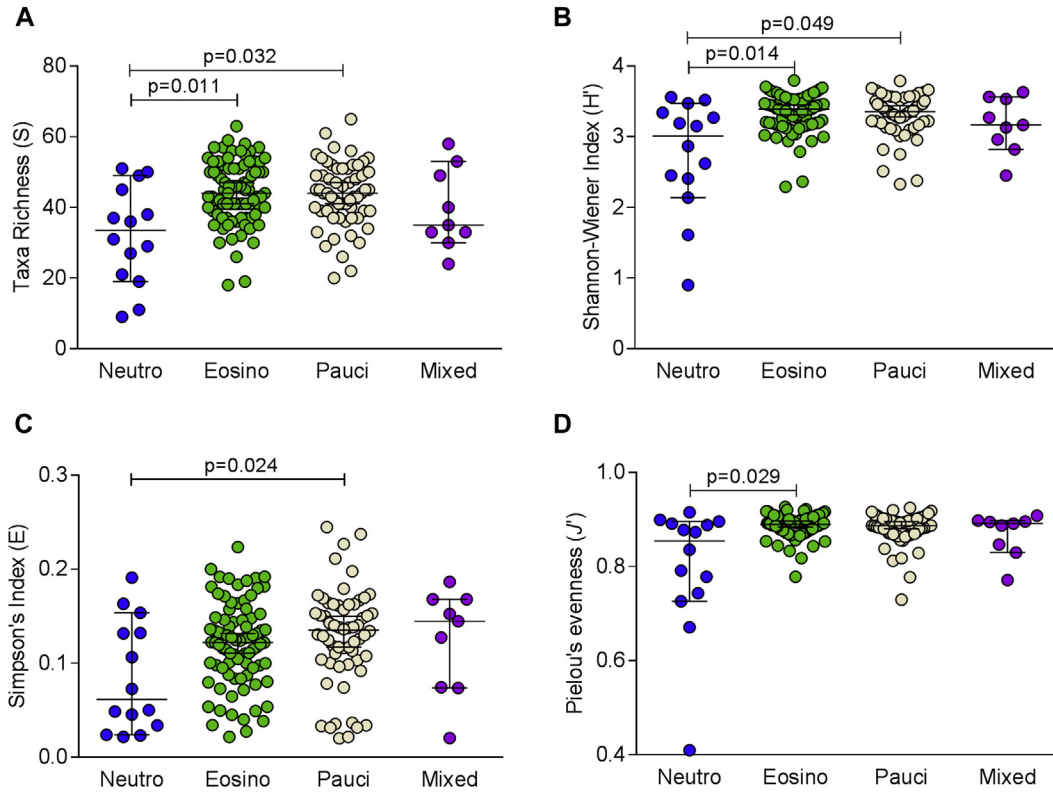
FIG E2. α-Diversity measures among asthma phenotypes: **A,** taxa richness; **B,** Shannon-Wiener index; **C,** Simpson evenness index; and **D,** Pielou evenness. *Bars* show medians ± 95% CIs. Statistical significance was assessed by using Kruskal-Wallis 1-way ANOVA with the Dunn *post hoc* test.
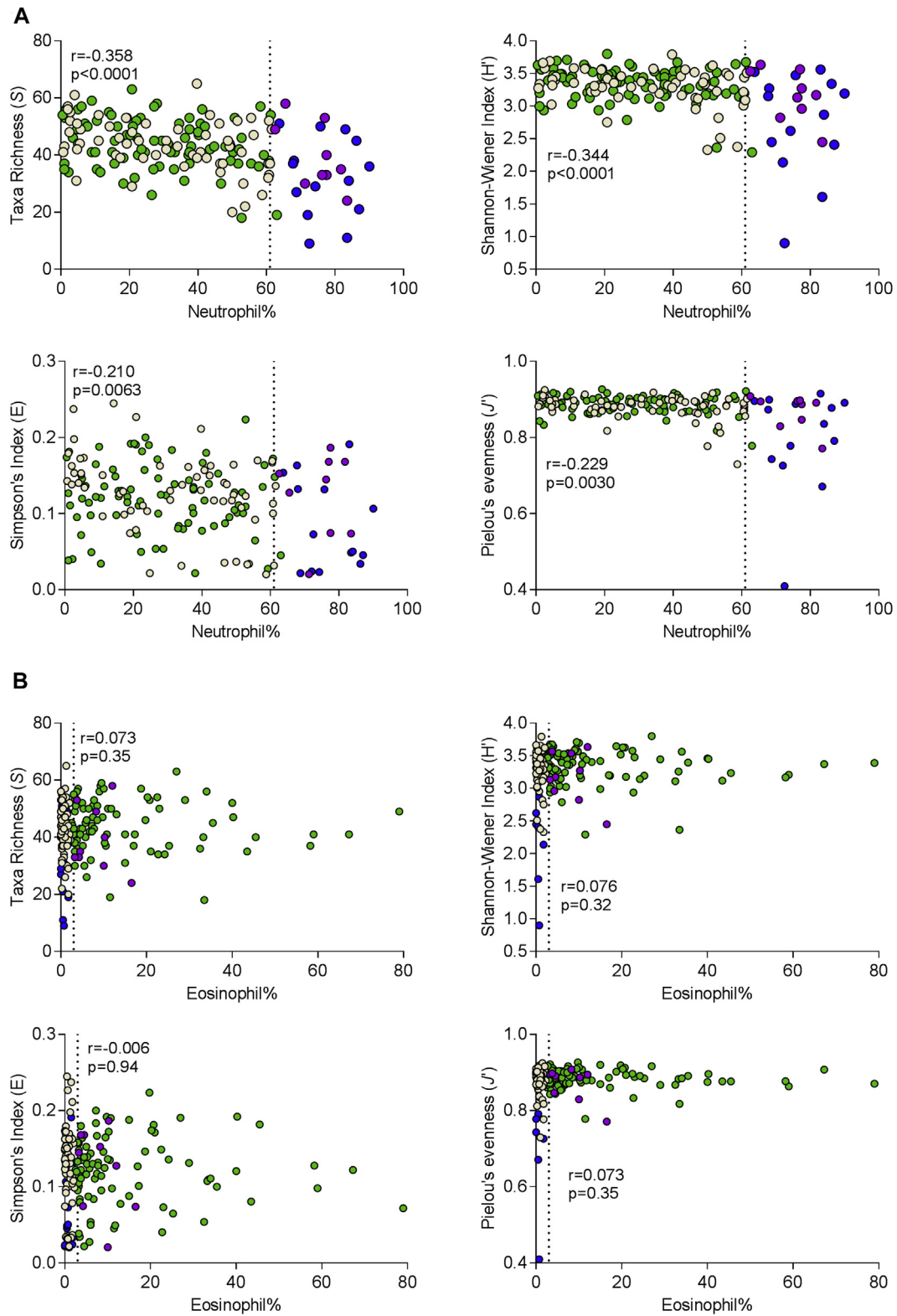
FIG E3. Correlations between sputum neutrophil/eosinophil counts (as a percentage of total cell count) and α-diversity measures. **A,** Neutrophil percentage, The *dotted line* at 61% neutrophils indicates phenotype cut-off point. **B,** Eosinophil percentage, The *dotted line* at 3% eosinophils indicates phenotype cutoff point. 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). Statistical significance was assessed by using Spearman rank correlation.
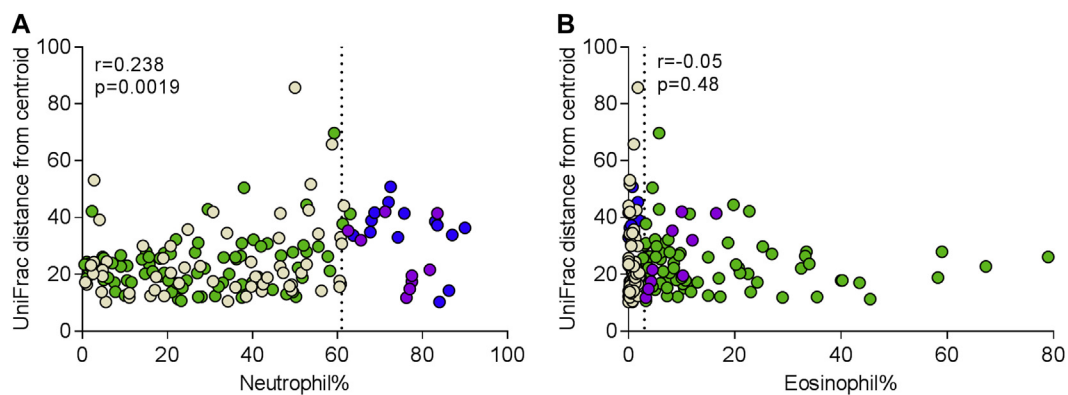
FIG E4. Correlations between sputum neutrophil/eosinophil counts (as a percentage of total cell count) and weighted UniFrac distance from centroid. **A,** Neutrophil percentage. The *dotted line* at 61% neutrophils indicates phenotype cutoff point. **B,** Eosinophil percentage. The *dotted line* at 3% eosinophils indicates phenotype cutoff point. 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). Statistical significance was assessed by using Spearman rank correlation.
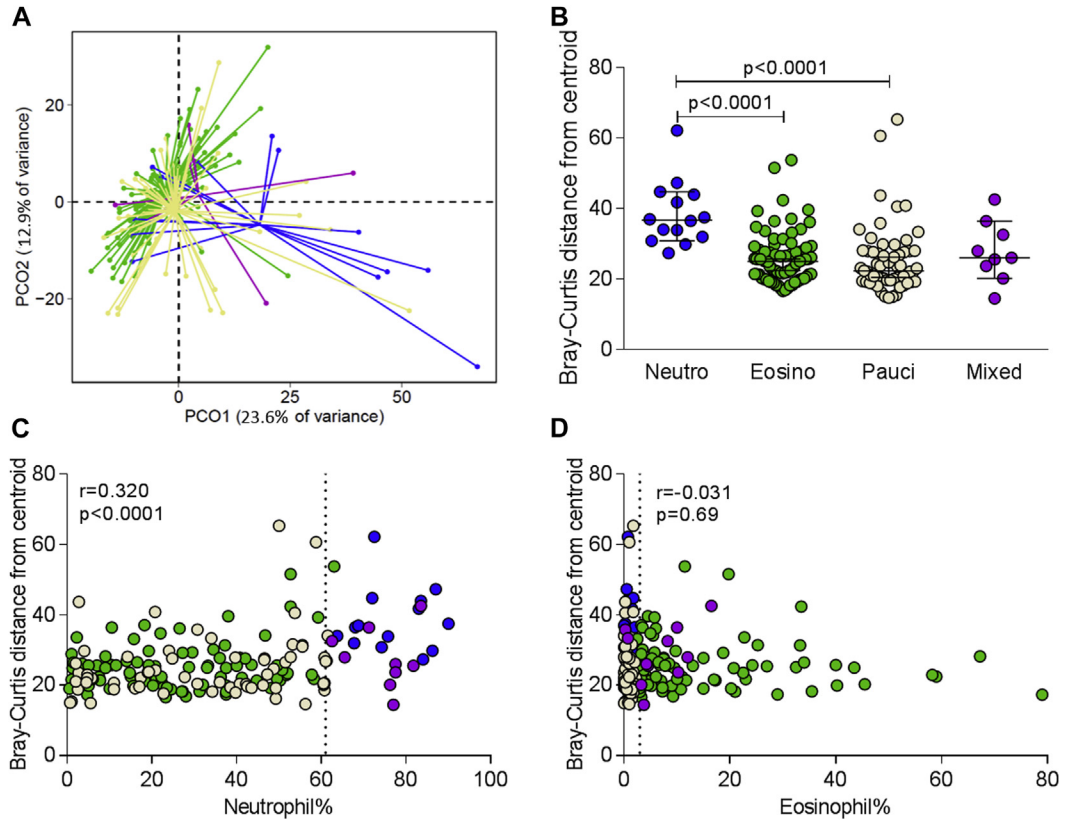
FIG E5. A, Principal coordinate analysis of asthma phenotype groups based on Bray-Curtis similarity distances. The first 2 principal coordinates are plotted on the *x*- and *y-axes*, respectively (representing 36.5% of total variation). **B,** Microbiota dispersion grouped by asthma phenotype. Distance from centroid was calculated from the Bray-Curtis dissimilarity matrix. **C** and **D,** Correlations between sputum inflammatory cell percentages and distance from centroid. Fig E5, *C*, Sputum neutrophil percentage versus Bray-Curtis distance from centroid. Fig E5, *D*, Sputum eosinophil percentage vs Bray-Curtis distance from centroid.

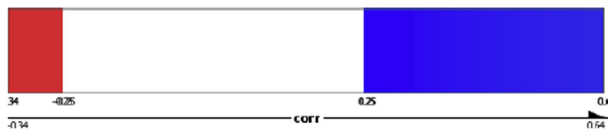**Node Fill Color Mapping**

subtraction

Node Size Mapping

Mean rel. ab.
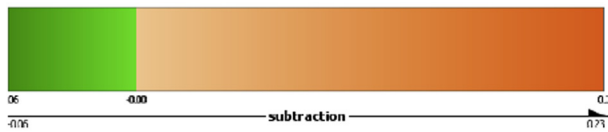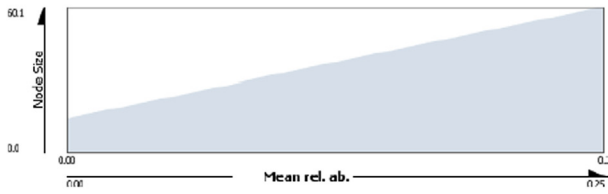
Edge Stroke Color (Unselected) Mapping
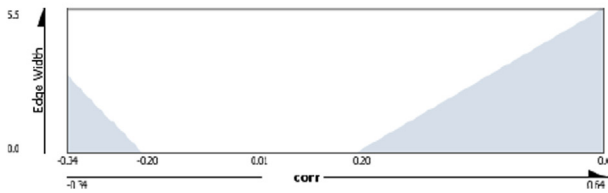
corr

Edge Width Mapping

corr

FIG E6. Bacterial network analysis, showing weight and color assigned to edges and nodes.
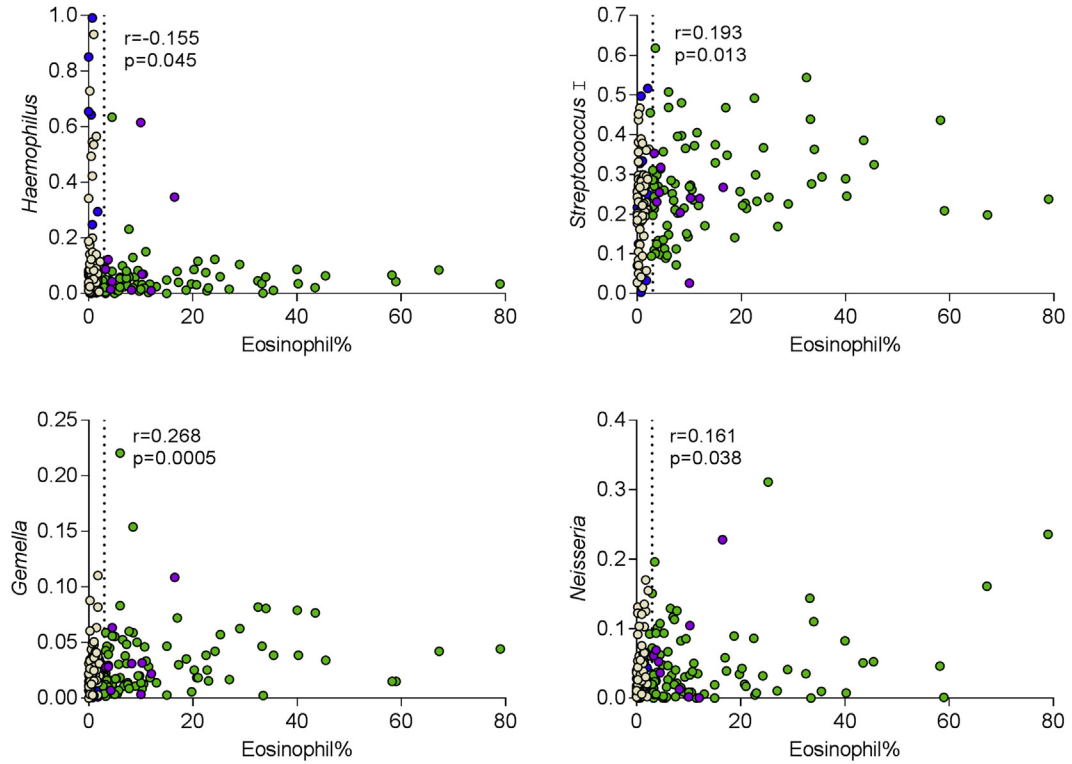
FIG E7. Taxa that significantly correlated with eosinophil percentages. Colors represent asthma phenotype based on neutrophilia or eosinophilia: blue is greater than 61% neutrophils, green is greater than 3% eosinophils, yellow is less than 61% neutrophils and less than 3% eosinophils (paucigranulocytic), and purple is both greater than 61% neutrophils and greater than 3% eosinophils (mixed). The *dotted line* at 3% eosinophils indicates phenotype cutoff points. Statistical significance was assessed by using Spearman rank correlation.
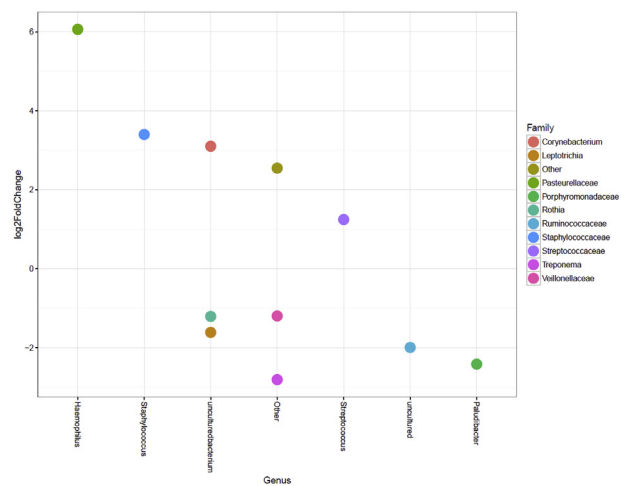
FIG E8. Normalized $\log_2$ fold changes of nonrarefied taxa read counts that significantly ($P < .05$) differed between neutrophilic and eosinophilic phenotypes. Positive fold change indicates taxa with significantly higher counts in neutrophilic participants, and negative fold change indicates taxa with significantly higher counts in eosinophilic participants. This shows that when *Haemophilus* or *Moraxella* taxa dominance do not influence data (because of nonrarefied count data as opposed to relative abundance), multiple taxa remain significantly different between neutrophilic and eosinophilic participants.

TABLE E1. 16S rRNA sequencing information

Median read count (Q1, Q3)	12,792 (8,060, 16,595)
Subsample depth	1,732
Samples excluded	7
Median Good coverage (Q1, Q3)	0.952 (0.942, 0.963)

Q1, Q3, Quartile 1, quartile 3.

TABLE E2. PERMANOVA analysis testing: significance of variance of weighted UniFrac and Bray-Curtis distance of sputum microbiota between asthma phenotypes (permutations = 9999)

Matrix	Source	df	SS	MS	Pseudo-F	P value (perm)
Weighted UniFrac	Phenotype	3	9,444.4	3,148.1	3.9969	.0004
	Residual	163	128,390	787.64		
	Total	166	137,830			
Bray-Curtis	Phenotype	3	8,215.4	2,738.5	3.3694	.0001
	Residual	163	132,480	812.74		
	Total	166	140,690			

TABLE E3. Pairwise PERMANOVA analysis testing: significance of variance of weighted UniFrac and Bray-Curtis distance of sputum microbiota between asthma phenotypes (permutations = 9999)

Matrix	Groups	T	P value (perm)	Unique perms
Weighted UniFrac	Pauci vs Neutro	2.52	<.01	9943
	Pauci vs Eosino	1.28	.13	9928
	Pauci vs Mixed	0.80	.62	9952
	Neutro vs Eosino	3.30	<.0001	9924
	Neutro vs Mixed	1.20	.21	9876
	Eosino vs Mixed	1.19	.19	9939
Bray-Curtis	Neutro vs Eosino	2.89	<.0001	9928
	Neutro vs Pauci	2.43	<.001	9918
	Neutro vs Mixed	1.18	.19	9888
	Eosino vs Pauci	1.15	.17	9921
	Eosino vs Mixed	0.96	.50	9899
	Pauci vs Mixed	0.78	.83	9922

Eosino, Eosinophilic; *Neutro*, neutrophilic; *Pauci*, paucigranulocytic.

TABLE E4. SIMPER analysis comparing taxa relative abundances between neutrophilic and eosinophilic phenotype groups and showing the 13 top contributing taxa, which collectively account for approximately 50% of variance between groups

Taxa	Neutrophilic Av. Abund	Eosinophilic Av. Abund	Av. Diss	Contrib %
Haemophilus	0.41	0.2	4.49	8.89
Prevotella	0.23	0.33	2.61	5.16
Streptococcus II	0.13	0.27	2.58	5.11
Streptococcus I	0.45	0.51	2.34	4.63
Veillonella	0.2	0.22	1.98	3.91
Moraxella	0.13	0	1.88	3.72
Neisseria	0.11	0.19	1.65	3.26
Rothia	0.1	0.19	1.52	3
Actinomyces sp. uncultured bacterium	0.1	0.16	1.5	2.96
Gemella	0.09	0.17	1.26	2.49
Leptotrichia	0.08	0.11	1.2	2.38
Actinomyces species oral clone DR002	0.05	0.08	1.1	2.17
Porphyromonas	0.04	0.11	1.1	2.17

TABLE E5. PERMANOVA analysis (top) and pairwise PERMANOVA on nondominant microbiome (battle), showing Bray-Curtis distance of sputum microbiota on the genera level grouped by asthma phenotype (permutations = 9999)

Source	df	SS	MS	Pseudo-F	P value (perm)
Phenotype	3	5,381	1,793.7	2.378	.0004
Residual	163	122,950	754.3		
Total	166	128,330			

Groups	T	P value (perm)	Unique perms
Neutro vs Eosino	2.31	<.0001	9922
Neutro vs Pauci	2.26	<.001	9911
Neutro vs Mixed	1.11	.24	9857
Eosino vs Pauci	0.91	.65	9904
Eosino vs Mixed	0.79	.86	9921
Pauci vs Mixed	0.73	.92	9924

Eosino, Eosinophilic; *Neutro*, neutrophilic; *Pauci*, paucigranulocytic.