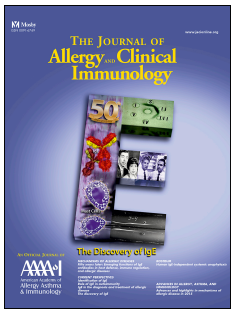
Accepted Manuscript



Inflammatory phenotypes in severe asthma are associated with distinct airway microbiology

Steven L. Taylor, BSc, Lex E.X. Leong, PhD, Jocelyn M. Choo, PhD, Steve Wesselingh, FRACP PhD, Ian A. Yang, FRACP PhD, John W. Upham, FRACP PhD, Paul N. Reynolds, MBBS PhD, Sandra Hodge, PhD, Alan L. James, FRACP PhD, Christine Jenkins, MBBS FRACP, Matthew J. Peters, MD FRACP, Melissa Baraket, PhD, Guy B. Marks, MBBS PhD, Peter G. Gibson, MBBS FRACP, Jodie L. Simpson, PhD, Geraint B. Rogers, PhD

PII: S0091-6749(17)30743-1

DOI: 10.1016/j.jaci.2017.03.044

Reference: YMAI 12790

To appear in: *Journal of Allergy and Clinical Immunology*

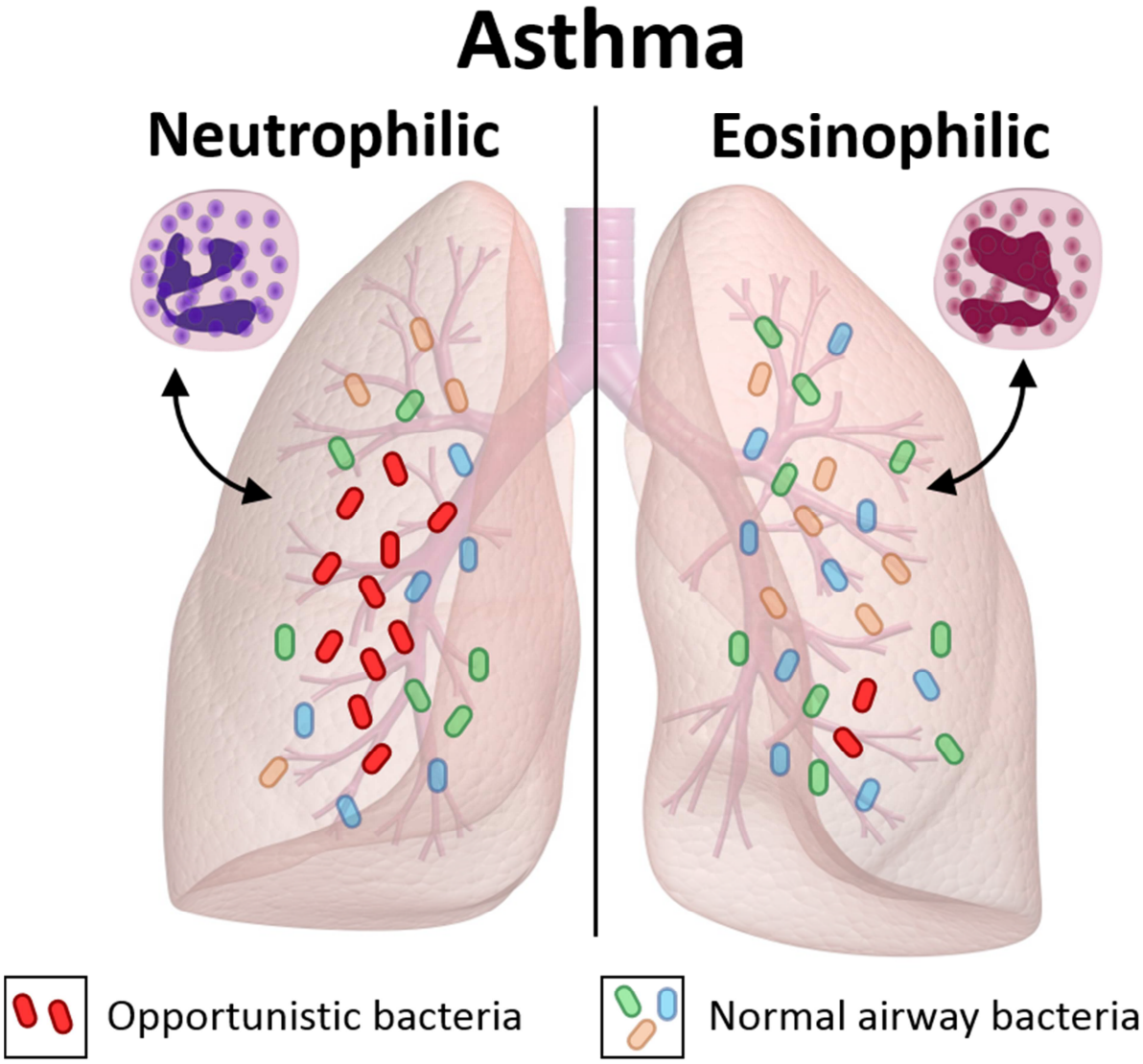
Received Date: 16 October 2016

Revised Date: 28 February 2017

Accepted Date: 15 March 2017

Please cite this article as: Taylor SL, Leong LEX, Choo JM, Wesselingh S, Yang IA, Upham JW, Reynolds PN, Hodge S, James AL, Jenkins C, Peters MJ, Baraket M, Marks GB, Gibson PG, Simpson JL, Rogers GB, Inflammatory phenotypes in severe asthma are associated with distinct airway microbiology, *Journal of Allergy and Clinical Immunology* (2017), doi: 10.1016/j.jaci.2017.03.044.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.


Asthma
Neutrophilic
Eosinophilic
Opportunistic bacteria
Normal airway bacteria

ACCEPTED

ACCEPTED MANUSCRIPT

Inflammatory phenotypes in severe asthma are associated with distinct airway microbiology

Steven L. Taylor BSc[1,2]*, Lex E.X. Leong PhD[1,2]*, Jocelyn M. Choo PhD[1,2], Steve Wesselingh FRACP PhD[1,2], Ian A. Yang FRACP PhD[3,4], John W. Upham FRACP PhD [3,5], Paul N. Reynolds MBBS PhD[6,7], Sandra Hodge PhD[6,7], Alan L. James FRACP PhD[8,9], Christine Jenkins MBBS FRACP[10,11], Matthew J. Peters MD FRACP[11,12], Melissa Baraket PhD[13,16], Guy B. Marks MBBS PhD[13,14,16], Peter G. Gibson MBBS FRACP[14,15], Jodie L. Simpson PhD[15†], Geraint B. Rogers PhD[1,2†]

[1] South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia

[2] SAHMRI Microbiome Research Laboratory, School of Medicine, Flinders University, Adelaide, South Australia, Australia

[3] School of Medicine, The University of Queensland, St Lucia, QLD, Australia.

[4] Department of Thoracic Medicine, The Prince Charles Hospital, Chermside, QLD, Australia.

[5] Translational Research Institute, Princess Alexandra Hospital, Woolloongabba, QLD, Australia.

[6] Department of Thoracic Medicine, Royal Adelaide Hospital and Lung Research Laboratory, Hanson Institute, Adelaide, SA, Australia.

[7] School of Medicine, The University of Adelaide, Adelaide, SA, Australia.

[8] Department of Pulmonary Physiology and Sleep Medicine, Sir Charles Gairdner Hospital, Nedlands, WA, Australia.

[9] School of Medicine and Pharmacology, University of Western Australia, Crawley, WA, Australia.

ACCEPTED MANUSCRIPT

[10] Respiratory Trials, The George Institute for Global Health, NSW, Australia,

[11] Australian School of Advanced Medicine, Macquarie University, NSW, Australia

[12] Department of Thoracic Medicine, Concord General Hospital, NSW, Australia,

[13] Respiratory Medicine Department and Ingham Institute, Liverpool Hospital, NSW, Australia

[14] Woolcock Institute of Medical Research, Glebe, NSW, Australia.

[15] Respiratory and Sleep Medicine, Priority Research Centre for Healthy Lungs, The University of Newcastle, Callaghan, NSW, Australia.

[16] South Western Sydney Clinical School, University of New South Wales, Sydney. NSW Australia

* Both authors contributed equally to this work

[†] Both authors contributed equally to this work

Corresponding author: Associate Professor Geraint Rogers, School of Medicine, Flinders University, University Drive, Bedford Park, Adelaide SA 5042, Australia.

Email: geraint.rogers@sahmri.com, Tel: +61 (0)8 8204 7614

Keywords: Asthma, microbiome, neutrophil, eosinophil

Declaration of interests

No author has a conflict of interest to declare in relation to manuscript.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Abstract

Background

Asthma pathophysiology and treatment responsiveness are predicted by inflammatory phenotype. However, the relationship between airway microbiology and asthma phenotype is poorly understood. We aimed to characterise 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 16S rRNA gene sequencing. Inflammatory phenotypes were defined by sputum neutrophil and eosinophil cell proportions. Microbiota were defined using alpha and beta diversity measures, and inter-phenotype differences identified using SIMPER, network analysis, and taxon fold change. Phenotypic predictors of airway microbiology were identified using multivariate linear regression.

Results

Microbiota composition was determined in 167 participants, classified as eosinophilic (n=84), neutrophilic (n=14), paucigranulocytic (n=60), or mixed neutrophilic-eosinophilic (n=9) phenotypes of asthma. Airway microbiology was significantly less diverse (p=0.022) and more dissimilar (p=0.005) in neutrophilic compared to eosinophilic participants. Sputum neutrophil proportion, but not eosinophil proportion, correlated significantly with these diversity measures (alpha-diversity: Spearman's r=-0.374, p<0.001; beta-diversity: r=0.238, p=0.002). Inter-phenotype differences were characterised by a greater frequency of pathogenic taxa at high relative abundance, and reduced *Streptococcus*, *Gemella* and

ACCEPTED MANUSCRIPT

Porphyromonas relative abundance in neutrophilic asthma. Multivariate regression confirmed sputum neutrophil proportion was the strongest predictor of microbiota composition.

Conclusions

Neutrophilic asthma is associated with airway microbiology that is significantly different to that in other inflammatory phenotypes, particularly eosinophilic asthma. Differences in microbiota composition may influence response to antimicrobial and steroid therapies, and risk of lung infection.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Key Messages

- Lower airway microbiology differs significantly between neutrophilic and eosinophilic patients with asthma, characterised by 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, gender, and atopy also independent predictors.
- The clear relationships between airway microbiota composition, inflammatory phenotype, and clinical measures suggest microbiota characterisation could be a useful contributor to the individualisation of asthma treatments.

Capsule summary

Utilizing a large, well-defined severe asthma cohort, we show that sputum neutrophilia, but not eosinophilia, was a strong predictor for airway microbiota diversity. These results highlight the divergence in pathophysiology based on asthma inflammatory phenotype.

Abbreviations.

ACQ6: Asthma Control Questionnaire 6

ANOVA: Analysis of variance

FEV_1: Forced expiratory volume in 1 second

FVC: Forced vital capacity

ICS: Inhaled corticosteroid

OTU: operational taxonomic unit

PERMANOVA: Permutational multiple ANOVA

SIMPER: Similarity of percentages

Role of Funding Source

ACCEPTED MANUSCRIPT

The research was supported by the Australian National Health and Medical Research Council (NHMRC) [grant number: 569246] and a grant from the John Hunter Hospital Charitable Trust.

ACCEPTED MANUSCRIPT

Introduction

Evidence-based patient stratification is required to address the heterogeneity of asthma severity and treatment responsiveness. Asthma phenotypes, based on characteristics of airway inflammation, are increasingly recognised as an important prognostic indicator.[1] 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 non-eosinophilic asthma).[2, 3] Indeed, based on relative numbers of sputum eosinophils and neutrophils, four inflammatory subtypes: eosinophilic asthma, neutrophilic asthma, mixed granulocytic asthma, and paucigranulocytic asthma, have been described.[2]

Unlike the relatively well-defined mechanisms that result in eosinophilic airway inflammation, those leading to non-eosinophilic asthma, particularly neutrophilic asthma, remain relatively poorly understood.[4] Further, while non-eosinophilic 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 microbiology. Compared with other patients with asthma, those with neutrophilic asthma are more likely to have a potentially pathogenic organism identified, by either culture-based[6] or culture-independent approaches,[5] and have reduced airway bacterial diversity.[7] Given that airway microbiota composition is associated with the degree of airway hyper-responsiveness among patients with sub-optimally controlled asthma,[8] inter-phenotype 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 bi-directional. Asthma phenotypes represent immunological 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.[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 asthma inflammatory phenotype. Defining the relationships between inflammatory phenotype and lower airway microbiota would inform our understanding of asthma pathophysiology, and could help to identify prognostic markers.

Several previous studies have reported differences in airway microbiology in eosinophilic and non-eosinophilic asthma, and the existence of significant relationships between this microbiota composition and clinical asthma measures.[5, 11-14] However, while providing important insight, these studies have involved relatively small and heterogeneous patient cohorts. Our study, based on participants enrolled into the Asthma and Macrolides: the Azithromycin Efficacy and Safety Study (AMAZES) trial (ACTRN12609000197235), was over three 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.

Through the application of a systematic approach to microbiota characterisation, 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 Supplementary Methods. Analysis was performed on samples collected as part of the baseline screening population from AMAZES clinical trial (ACTRN12609000197235). Participants were recruited from eight centres across Australia (detailed in Supplementary Methods). Pre-defined inflammatory phenotype categories, based on sputum cell counts relative to patient age, were assigned as published previously and detailed in the Supplementary Methods.[15, 16] Briefly, neutrophilic phenotype was defined as ≥61% neutrophils (neutrophil% cut-off dependent on age), eosinophilic phenotype as ≥3% eosinophils, paucigranulocytic as ≤61% neutrophils and ≤3% eosinophils and mixed granulocytic was defined as ≥61% neutrophils and ≥3% eosinophils.

DNA extraction and 16S rRNA gene amplicon sequencing and gene copy number

DNA extraction was performed on 100 µl sputum aliquots using a combined physical, enzymatic, and heat-based cell lysis, followed by phenol-chloroform extraction and DNA recovery using EZ-10 Spin columns (Bio Basic, Inc., Ontario, Canada). V1-3 hypervariable region of the bacterial 16S rRNA gene was amplified from sputum DNA using modified primers 27F and 519R. Amplicons were cleaned, indexed and sequenced according to the Illumina MiSeq 16S Metagenomic Sequencing Library Preparation protocol. 16S rRNA sequence data was processed as previously described.[17] Spurious operational taxonomic units (OTUs) were removed systematically using previous reports of common laboratory sequencing contaminants.[18] Bacterial burden was quantified using qPCR for the 16S rRNA gene.[19] Detailed extraction, sequencing and qPCR protocols are provided as Supplementary Methods.

Diversity measurements and statistical analyses

Five alpha diversity (within-sample variance) indices were employed to test a variety of parameters of within-patient taxon distribution: Faith's phylogenetic diversity (where a higher value indicates a more phylogenetically diverse sample), Simpson's and Pielou's evenness indices (where 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). Beta diversity (inter-sample variance) was determined by using two approaches: weighted UniFrac similarity (which accounts for phylogenetic distance) and square-root transformed Bray-Curtis similarity (based on the relative abundance of taxa alone). Alpha and beta diversity measures were calculated using either QIIME (v. 1.8.0) or Primer (v. 6, PRIMER E Ltd, Plymouth, UK).

Continuous data were tested for non-normality including skewness and kurtosis using the D'Agostino-Pearson omnibus test. The Kruskal-Wallis one-way ANOVA with Dunn's post hoc test was used for multiple comparisons of non-normally distributed data, Mann-Whitney U test for pairwise comparisons, chi-squared test for categorical data, and Spearman's test for correlations (GraphPad PRISM; v.7.01 GraphPad Software Inc., California, USA).

Multivariate linear regression was performed using Faith's phylogenetic diversity and UniFrac distance from centroid as two dependent variables reflecting aspects of diversity (SPSS v.23.0, IBM, Armonk, NY). Covariates were selected a priori and included in the model based on a significant correlation with either dependent variable. Confidence intervals were obtained by bootstrapping; resampling 1,000 times. Covariates were tested for collinearity using variance inflation factors.

ACCEPTED MANUSCRIPT

154 *Taxon dispersion*

155 Variation in microbiota composition between groups was assessed using SIMilarity of
156 PERcentages (SIMPER) analysis in PRIMER, and SparCC from which correlations ($r \geq 0.25$
157 or $r \leq -0.25$) and p values ($p \leq 0.01$) were visualized by network analysis using Cytoscape
158 (v.3.4.0). Pathogen overgrowth adjustment is described in Supplementary Methods.

ACCEPTED MANUSCRIPT

Results

Clinical characteristics

Induced sputum samples were obtained from 187 participants. Of these, thirteen were excluded due to poor sample quality. Of the 174 that underwent 16S rRNA gene amplicon sequencing, a further seven were excluded due to an insufficient sequence read depth (Supplementary Table 1). The remaining 167 subjects were classified as one of four inflammatory phenotypes: neutrophilic (n=14), eosinophilic (n=84), paucigranulocytic (n=60), or mixed granulocytic (n=9) based on previously described sputum inflammatory cell count percentages.[15, 16] There was no significant difference in age, gender distribution, atopy, smoking history, ICS dose, GINA treatment step, or mean ACQ6 score between these phenotypic groups, as assessed using multiple comparison tests (Table 1). There were, however, significant differences in lung function, as assessed by both FEV_1 % predicted (p=0.035) and FEV_1/FVC% (p=0.013).

Following quality filtering and chimera removal, 16S rRNA gene amplicon sequencing resulted in a median read depth of 12,792 (q1, q3: 8060, 16595). Sequence data was sub-sampled 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=0.51, Kruskal-Wallis test, Supplementary Figure 1).

Alpha diversity

Participants with neutrophilic asthma had significantly lower Faith's phylogenetic score (p=0.022) than participants with eosinophilic asthma, which resembled that of paucigranulocytic asthma (Figure 1A). Faith's phylogenetic diversity significantly correlated with the sputum neutrophil % (r=-0.374, p<0.0001; Figure 1B) but not with sputum eosinophil % (r=0.146, p=0.060; Figure 1C). Analysis with a range of alternative alpha

diversity indices (taxa richness, Shannon-Wiener index, Simpson's index, and Pielou's evenness; Supplementary Figures 2 and 3), resulted in consistent findings in relation to phenotype, sputum neutrophil % and sputum eosinophil %. Together, these results demonstrate a significant relationship between airway microbiota composition and sputum neutrophilia, but not sputum eosinophilia.

Beta diversity

Principal Coordinate Analysis (PCoA) of weighted UniFrac similarity distance showed that neutrophilic samples were distinguished from other phenotypes along the first and second principal coordinates, while the other phenotypes broadly clustered together (Figure 2A). Consistent with these observations, a PERMANOVA test showed that phenotype grouping contributed significantly to the differences in microbial composition of the samples ($p=0.0004$, pseudo-$F=3.997$, Supplementary Table 2). Pairwise PERMANOVA comparing the phenotype groups indicated that variance was attributed to neutrophilic vs eosinophilic ($p=0.0001$, $T=3.30$) and neutrophilic vs paucigranulocytic ($p=0.0015$, $T=2.52$) groups (Supplementary Table 3).

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 (Figure 2B). In keeping with alpha diversity analyses, variance in distance from centroid was associated with sputum neutrophil %, rather than sputum eosinophil % (Supplementary Figure 4). Beta-diversity analyses using a second distance measure, Bray-Curtis, gave consistent findings (Supplementary Figure 5).

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 (Supplementary Table 4). Hierarchical cluster analysis based on relative taxon abundance revealed that *Moraxella* and *Haemophilus* clustered separately from the other eleven taxa (Figure 3). In the patients with neutrophilic asthma, *Moraxella* and *Haemophilus* exceeded 40% relative abundance in 6/14 (42.9%) samples compared to only 1/84 (1.19%) patients with eosinophilic asthma, 7/60 (11.7%) with paucigranulocytic and 1/9 (11.1%) with mixed (χ^2=25.5, p<0.0001, Figure 3). Relationships between bacterial taxon relative abundance were further visualised by network analysis (Figure 4, Supplementary Figure 6), revealing a bacterial community of taxa whose abundance is positively correlated in almost all cases. Most of these taxa were more prevalent in eosinophilic samples than in neutrophilic samples. *Haemophilus*, 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 Supplementary Methods for classification), *Gemella*, *Rothia* and *Porphyromonas* were significantly less abundant in neutrophilic than in eosinophilic and paucigranulocytic phenotypes (Figure 5A). Sputum neutrophil % positively correlated with the relative abundance of *Moraxella* and negatively correlated with the relative abundance of *Streptococcus* I, *Gemella* and *Porphyromonas* (Figure 5B). In contrast, *Haemophilus* negatively correlated with eosinophil %, and *Streptococcus* I, *Neisseria* and *Gemella* positively correlated with eosinophil % (Supplementary Figure 7). *Prevotella, Actinomyces, Leptotrichia, and Veillonella*, while

identified by SIMPER and represented highly connected nodes in the network analysis, did not differ between airway inflammatory phenotypes, nor correlate with sputum cell counts.

Non-dominant microbiome

We sought to establish whether differences in microbiota composition between inflammatory phenotypes were explained solely by overgrowth of opportunistic taxa (e.g. *Haemophilus* and *Moraxella*) or, alternatively, whether differences existed even in the absence of pathogen predominance. Two separate approaches were used to investigate this; re-scaling of relative abundance data, following the exclusion of pathogen predominance, and assessment of ranked taxon fold change between neutrophilic and eosinophilic groups based on non-subsampled taxa counts. Re-scaled relative abundance data remained significantly different between inflammatory phenotypes (p=0.0004, pseudo-F=2.38, Supplementary Table 5). Pairwise tests revealed significant differences between neutrophilic and eosinophilic phenotype participants (p=0.0001, T=2.31) and between neutrophilic and paucigranulocytic participants (p=0.0002, T=2.26, Supplementary Table 5). Assessment of taxa count ranked fold change supported these findings, with significant taxa count differences between participants with neutrophilic and eosinophilic phenotypes (Supplementary Figure 8).

Clinical and inflammatory associations with microbiota composition

In univariate analysis, Faith's phylogenetic diversity significantly, inversely correlated with sputum neutrophil %, age, and ICS dose, and significantly, positively with FEV_1 % (Table 2). Conversely, weighted UniFrac distance from centroid significantly correlated, positively with sputum neutrophil %, and was significantly different based on gender and atopy but not with age, ICS dose, FEV_1 % or previous smoking status (Table 2). In multivariate analysis sputum neutrophil % was the only variable that independently predicted both Faith's phylogenetic

ACCEPTED MANUSCRIPT

diversity and weighted UniFrac distance from centroid (p=0.002, 95% CI=-0.07, -0.02 and p<0.001, 95% CI=0.07, 0.22, respectively; Table 3). Age and ICS dose both independently predicted Faith's diversity (p=0.030, 95% CI=-0.07, -0.004 and p=0.042, 95% CI=-0.001, -0.001, respectively) while atopy and gender independently predicted distance from centroid (p=0.018, 95% CI=1.3, 8.4 and p=0.039, 95% CI=-7.5, -0.30, respectively).

ACCEPTED MANUSCRIPT

Discussion

To our knowledge, this is the largest study to date to assess predictors of the airway microbiota composition in asthma. Our primary comparisons were between asthma inflammatory phenotypes, where we observed significant differences in the composition of airway microbiota. These differences were largely between neutrophilic and eosinophilic participants, and reflected a reduced diversity and evenness of detectable bacterial taxa in the neutrophilic participants. Reduced microbiota diversity has been reported following acute and chronic airway infections in asthma,[5, 7] and in other respiratory disorders,[20-22] as well as with the effects of exposure to antibiotics.[23, 24] Importantly, none of the study participants reported clinical features of respiratory infection or had antibiotic therapy during the preceding month.

We further assessed alpha diversity metrics relative to continuous neutrophil and eosinophil count data as an alternative to categorical inflammatory phenotypes. Significant correlations were observed between sputum neutrophil % and each assessed alpha diversity metric, with no significant interactions between any diversity metric and sputum eosinophil %, strongly suggesting that decreased microbiota richness, evenness, and diversity are associated with airway neutrophilia. Analysis of sputum microbiota beta diversity (inter-sample similarity) also demonstrated substantial differences between people with neutrophilic airway inflammation and those with other inflammatory phenotypes, consistent with the stochastic overgrowth of complex commensal communities by individual opportunistic pathogens.[25]

We identified bacterial taxa that contributed to observed differences in microbiota composition between inflammatory phenotypes, namely high abundance of *Haemophilus* and *Moraxella* 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 sub-clinical 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 % (*Gemella*, *Porphyromonas*, and *Streptococcus*) and, importantly, even after overgrowth effects of *Haemophilus* and *Moraxella* were controlled for, significant differences in the microbiota composition between neutrophilic and eosinophilic participants were still observed. This finding suggests that two separate phenomena contribute to microbial differences between inflammatory phenotypes; the impact 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 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, where low-dose macrolide and ICS therapies have been shown to influence overgrowth by opportunistic respiratory pathogens and innate immune function, respectively.[6, 14, 26] Further, the relative lack of efficacy of ICS in patients with non-eosinophilic asthma[2] may lead to use of higher doses compared to 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]

While strong associations between 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* 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.[27]

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 characterisation from these participants then allowed us to assess the extent to which clinical and inflammatory characteristics independently associated with variations in airway microbiology using multivariate linear regression analysis. Multivariate regression identified sputum neutrophil % as the strongest predictor of microbiota variance. However age, ICS dose, gender and atopy were also significant, independent predictors. Conversely, lung function (as measured by FEV_1 %) and smoking status were not. Of particular interest was the finding that increasing age predicted reduced microbiota alpha diversity, as age has been previously associated with microbiome composition in other chronic respiratory diseases. The airway microbiome of cystic fibrosis patients is strongly affected by age,[28] 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 may reflect the effect of age on neutrophilia,[15] suggesting that a tendency towards a neutrophilic phenotype and/or a susceptibility to opportunistic airway infections increases with age in patients with severe asthma.

It is important to recognise a number of limitations of our study. Airway microbiology was assessed based on induced sputum, which, while shown to provide reproducible inflammatory cell levels in patients with moderate to severe asthma,[29, 30] only provides an approximation of lower airway microbiology (in common with other lower airway sampling strategies).[31] It is also important to note that induced sputum levels of neutrophils and eosinophils may change frequently,[32, 33] 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

was subsampled to a level that allowed the inclusion of the greatest number of subjects, while maintaining sufficient read depth to accurately describe microbiota composition. Additional analysis, using a greater read depth, might however identify rare taxa that contribute to disease characteristics. While multivariate regression identified sputum neutrophilia as an independent predictor of microbiota composition, the effects of variation in ICS dose between phenotypes (although non-significant) should be noted. Finally, though none of the study participants received antibiotics in the month prior to recruitment, data on less recent exposure were not available and could have a lasting impact on the lower airway microbiome composition.[23]

The clear relationship between airway inflammatory phenotype and microbiota highlight the need for studies examining whether asthma treatments should be individualised 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 characterisation could be used as a basis for asthma treatment selection.

Acknowledgements

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, I-Chin Wu, Monique de Pedro, Kirsty Herewane, Miranda Ween, Fungai Madzinga, Melissa McClean, Pamela Fung, Janet Shaw, Joanne McNamara, , Kevin Oreo, Peta Grayson, Robyn Jones, Tessa 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.

References

1. Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nature medicine* 2012; **18**(5): 716-25.
2. Pavord ID, Brightling CE, Woltmann G, Wardlaw AJ. Non-eosinophilic corticosteroid unresponsive asthma. *Lancet* 1999; **353**(9171): 2213-4.
3. Simpson JL, Scott R, Boyle MJ, Gibson PG. Inflammatory subtypes in asthma: assessment and identification using induced sputum. *Respirology* 2006; **11**(1): 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. *Mediators of inflammation* 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**(6): 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**(7): 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. *The European respiratory journal* 2016; **47**(3): 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. *The Journal of allergy and clinical immunology* 2011; **127**(2): 372-81 e1-3.
9. Rogers GB, Hoffman LR, Carroll MP, Bruce KD. Interpreting infective microbiota: the importance of an ecological perspective. *Trends in microbiology* 2013; **21**(6): 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**(6): 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. *The Journal of allergy and clinical immunology* 2015; **136**(4): 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**(3): 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**(4): 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. *The Journal of allergy and clinical immunology* 2016; **137**(5): 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**(5): 857-65.
16. Simpson JL, Scott R, Boyle MJ, Gibson PG. Inflammatory subtypes in asthma: assessment and identification using induced sputum. *Respirology* 2006; **11**(1): 54-61.
17. 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.
18. 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 Biology* 2014; **12**(1): 1-12.

19. 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**(1): 257-66.
20. 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. *The ISME journal* 2016; **10**(5): 1081-91.
21. 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 infectious diseases* 2015; **15**: 336.
22. 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 infectious diseases* 2016; **16**: 317.
23. 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. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 2013; **12**(1): 22-8.
24. Huang YJ, Sethi S, Murphy T, Nariya S, Boushey HA, Lynch SV. Airway microbiome dynamics in exacerbations of chronic obstructive pulmonary disease. *Journal of clinical microbiology* 2014; **52**(8): 2813-23.
25. 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.
26. 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. *The Lancet Respiratory medicine* 2014; **2**(12): 988-96.
27. 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. *The Journal of allergy and clinical immunology* 2016.
28. Zhao J, Schloss PD, Kalikin LM, Carmody LA, Foster BK, Petrosino JF, et al. Decade-long bacterial community dynamics in cystic fibrosis airways. *Proceedings of the National Academy of Sciences of the United States of America* 2012; **109**(15): 5809-14.
29. 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. *Mediators of inflammation* 2002; **11**(5): 293-8.
30. Rossall MR, Cadden PA, Molphy SD, Plumb J, Singh D. Repeatability of induced sputum measurements in moderate to severe asthma. *Respiratory medicine* 2014; **108**(10): 1566-8.
31. 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. *Journal of clinical microbiology* 2010; **48**(1): 78-86.
32. 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**(3): 461-6.
33. 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. *Archivos de bronconeumologia* 2016; **52**(2): 76-81.

Table 1: Clinical and inflammatory cell parameters of participants

	Neutrophilic	Eosinophilic	Paucigranulocytic	Mixed Granulocytic	p
N	14	84	60	9	
Age, mean (SD)	59.8 (13.9)	57.2 (15.2)	55.8 (13.8)	56.8 (17.8)	0.817
Male, n (%)	9 (64.3)	30 (35.7)	23 (38.3)	6 (66.7)	0.084
Atopic, n (%)	11 (78.6)	62 (76.5), n=81	52 (86.7)	6 (66.7)	0.305
Previous smoker, n (%)	6 (42.9)	32 (38.1)	17 (28.3)	4 (44.4)	0.491
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)	0.379
Duration of asthma, years 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)	0.205
FEV_1 % predicted, mean (SD)	70.3 (18.2)	70.0 (17.9)	78.7 (19.2)*	68.4 (17.3)	0.035
FVC % predicted, mean (SD)	81.4 (11.6)	83.6 (16.4)	85.1 (15.0)	82.9 (11.7)	0.854
FEV_1/FVC %, mean (SD)	65.8 (14.5)	64.9 (12.2)	71.3 (12.2)*	63.9 (12.7)	0.013
ACQ6, mean (SD)	2.1 (1.2)	1.9 (0.9)	1.6 (0.8)	1.5 (0.6)	0.210
GINA Treatment step 1 2 3 4 5	 1 (7.1) 0 (0) 0 (0) 11 (78.6) 2 (14.3)	 2 (2.5) 0 (0) 10 (12.4) 67 (82.7) 2 (2.5)	 0 (0) 0 (0) 12 (20.3) 47 (80.0) 0 (0)	 0 (0) 0 (0) 2 (22.2) 7 (77.8) 0 (0)	0.058
ICS dose, µg, median (q1,q3)	2000 (1280,2000), n=13	1000 (800,2000), n=82	1000 (800,2000), n=59	1600 (1000, 2000)	0.254
Total cell count, x 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)	<0.001
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)* †	<0.001
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)* †	<0.001
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)# †	<0.001
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) *†	<0.001
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)* †	0.018

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) †	<0.001
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)	0.002

ACQ6: asthma control questionnaire 6 question format, FEV_1: forced expiratory volume in 1 second, FVC: forced vital capacity, ICS: inhaled corticosteroids, SD: standard deviation, q1, q3: quartile 1, quartile 3. P values in last column describe variance across four phenotypes. *p<0.05 vs eosinophilic asthma, †p<0.05 vs. paucigranulocytic, #p<0.05 vs. neutrophilic asthma.

ACCEPTED MANUSCRIPT

Table 2: Comparison of patient characteristics with alpha diversity (Faith's phylogenetic diversity) and beta diversity (weighted UniFrac distance from centroid). Assessed using Spearman or Mann-Whitney test.

		Neutrophil %	Age	ICS Dose*	FEV_1 % predicted	Atopy†#	Gender†	Ever smoked†
Faith's diversity	r	-0.374	-0.309	-0.242	0.193			
	p	<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	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
#n=164
†Assessed using Mann-Whitney test

Table 3: Multivariate linear regression on alpha diversity (Faith's phylogenetic diversity) and beta diversity (weighted UniFrac distance from centroid). Performed on 160 participants.

		B	95% CI	p
Faith's phylogenetic diversity	Neutrophil %	-0.046	-0.07, -0.02	0.002
	Age	-0.036	-0.07, -0.004	0.030
	Gender	0.35	-0.59, 1.4	0.49
	Atopy	-0.77	-1.7, 0.12	0.10
	FEV_1 % predicted	0.016	-0.01, 0.04	0.23
	ICS dose	-0.001	-0.001, -0.001	0.042
UniFrac distance from centroid	Neutrophil %	0.14	0.07, 0.22	<0.001
	Age	0.056	-0.05, 0.16	0.30
	Gender	-3.9	-7.5, -0.30	0.039
	Atopy	4.8	1.3, 8.4	0.018
	FEV_1 % predicted	0.049	-0.04, 0.14	0.27
	ICS dose	<0.001	-0.001, 0.001	0.56

ACCEPTED MANUSCRIPT

Figure Legends

Figure 1: Faith's phylogenetic diversity is significantly associated with sputum neutrophilia but not eosinophilia. A) Patients grouped by asthma phenotype. B) Neutrophil % where dotted line at 61% neutrophils indicates phenotype cut-off point. C) Eosinophil % where dotted line at 3% eosinophils indicates phenotype cut-off point. Colours represent asthma phenotype, where blue= >61% neutrophils, green= >3% eosinophils, yellow= <61% neutrophils and <3% eosinophils (paucigranulocytic), and purple= both >61% neutrophils and >3% eosinophils (mixed). Statistical significance was assessed by A) Kruskal-Wallis one-way ANOVA with Dunn's post hoc test or B) and C) Spearman's rank correlation.

Figure 2: Microbiota dispersion grouped by asthma phenotype. A) Principal Coordinate Analysis (PCoA). The first two 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 Kruskal-Wallis one-way ANOVA with Dunn's post hoc test.

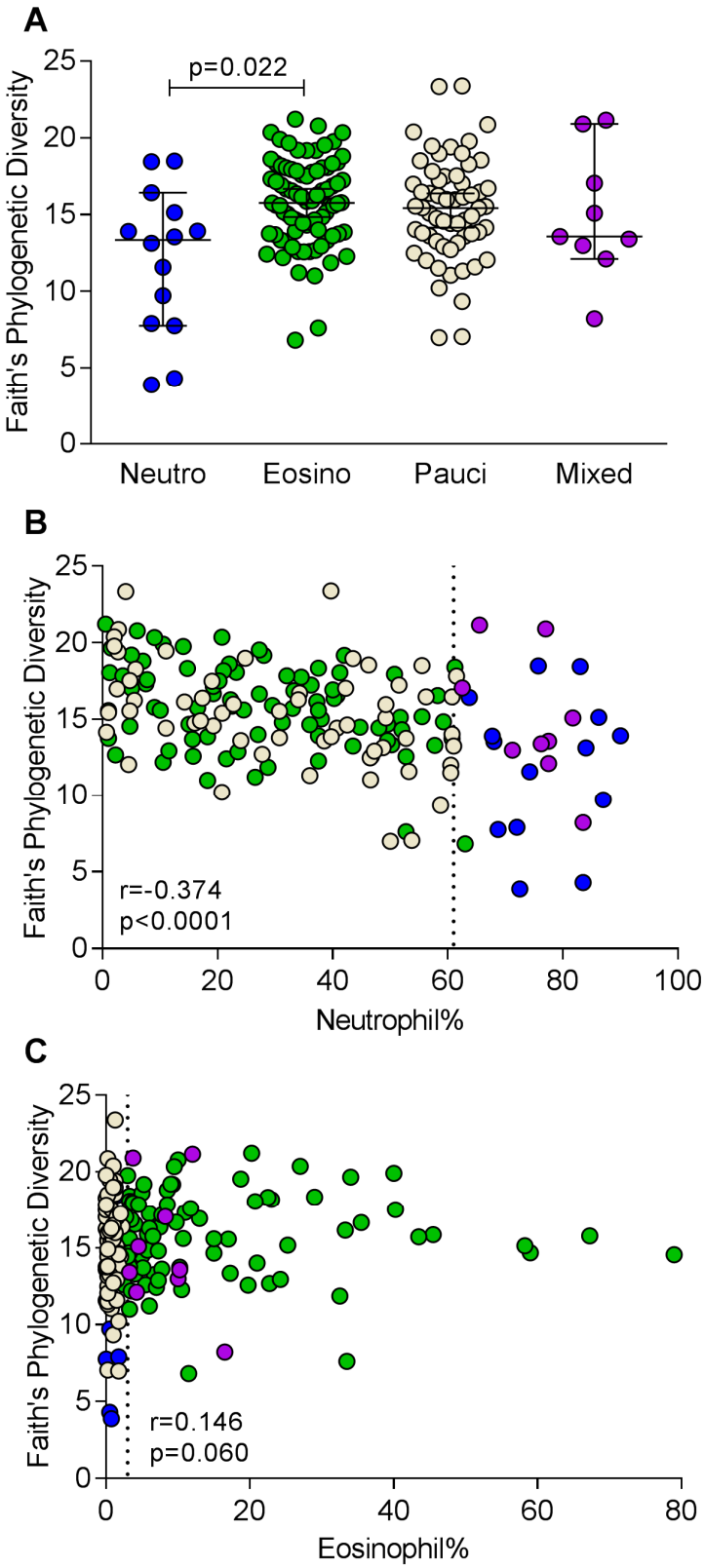
Figure 3: Relative abundance of discriminant taxa among asthma phenotypes. The thirteen taxa that collectively contribute to approximately 50% of variance among phenotypes, as determined by SIMPER analysis. The clustering shows the similarity relationship of genera based on Bray-Curtis similarity distance and single linkage hierarchical clustering method. * *Actinomyces* sp. uncultured bacteria, # *Actinomyces* sp. oral clone DR002.

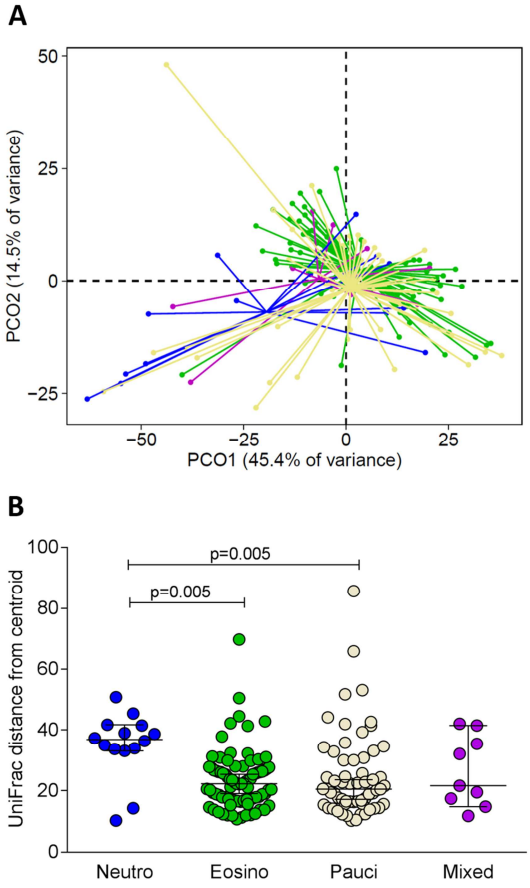
Figure 4: Bacterial network analysis of asthma cohort. Each edge represents a significant correlation coloured by either positive (blue) or negative (red). Edge width and transparency are proportional to the absolute value of the correlation coefficient. Node size is proportional to mean relative abundance. Node hue is proportional to the difference in taxon relative

abundance between the neutrophilic phenotype group and the eosinophilic phenotype group. Correlations performed by SparCC with a correlation cut-off of R>0.25 or <-0.25. * *Actinomyces* sp. uncultured bacteria, # *Actinomyces* sp. oral clone DR002.

Figure 5: Taxa distribution differs by sputum neutrophilia. A) Taxa which significantly differ by patient inflammatory phenotype. B) Significant correlations between taxa and neutrophil %. Colours represent asthma phenotype, based on neutrophilia or eosinophilia where blue= >61% neutrophils, green= >3% eosinophils, yellow= <61% neutrophils and <3% eosinophils (paucigranulocytic), and purple= both >61% neutrophils and >3% eosinophils (mixed). Dotted line at 61% neutrophils indicates phenotype cut-off point. Statistical significance was assessed by A) Kruskal-Wallis one-way ANOVA with Dunn's post hoc test and B) Spearman's rank correlation.

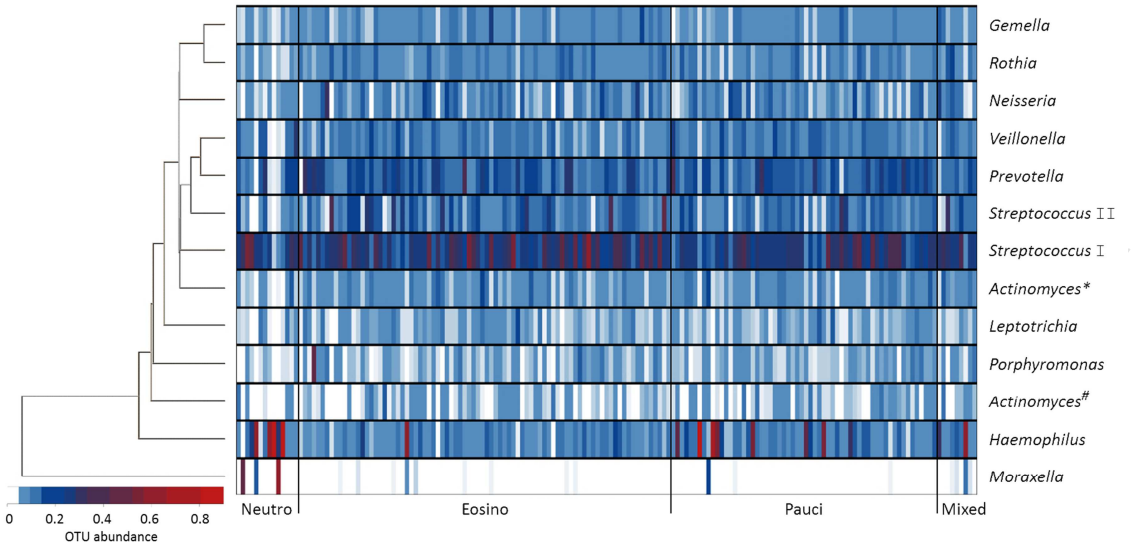
ACCEPTED MANUSCRIPT


A
p=0.022
Faith's Phylogenetic Diversity
25
20
15
10
5
0
Neutro
Eosino
Pauci
Mixed
B
Faith's Phylogenetic Diversity
r=-0.374
p<0.0001
0
20
40
60
80
100
Neutrophil%
C
Faith's Phylogenetic Diversity
r=0.146
p=0.060
0
20
40
60
80
Eosinophil%

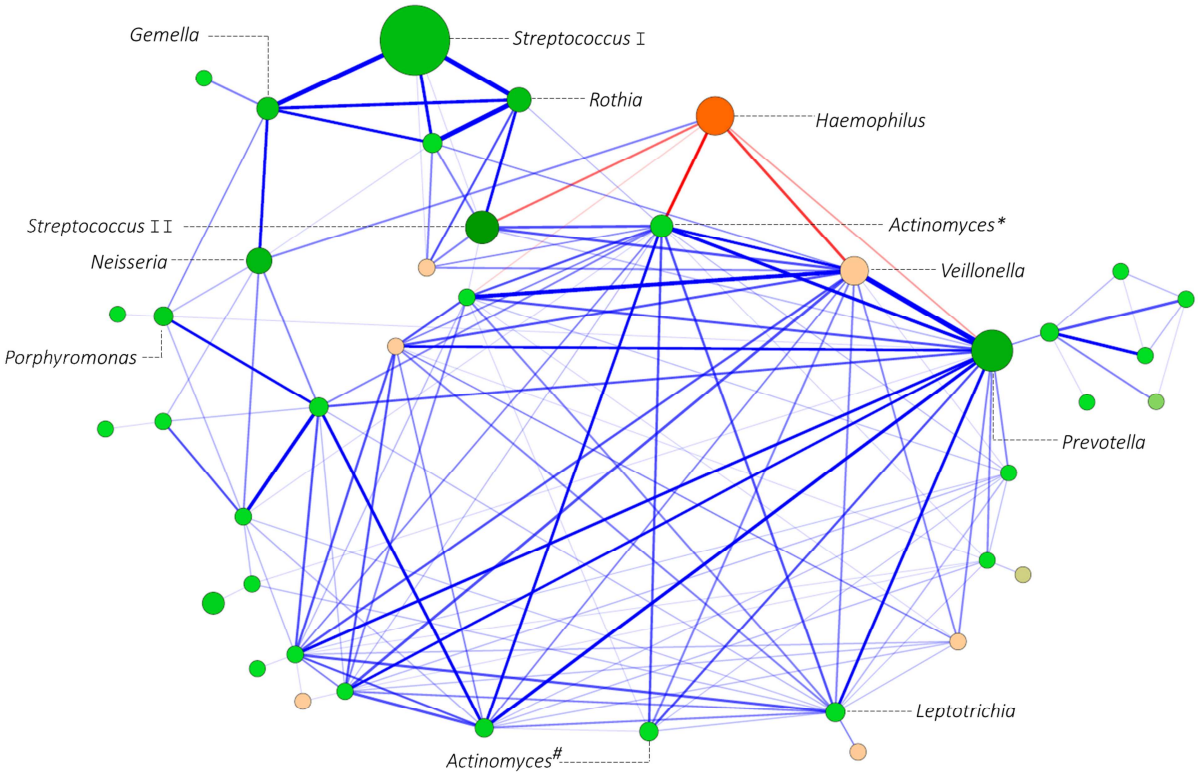


ACCEPTED MANUSCRIPT

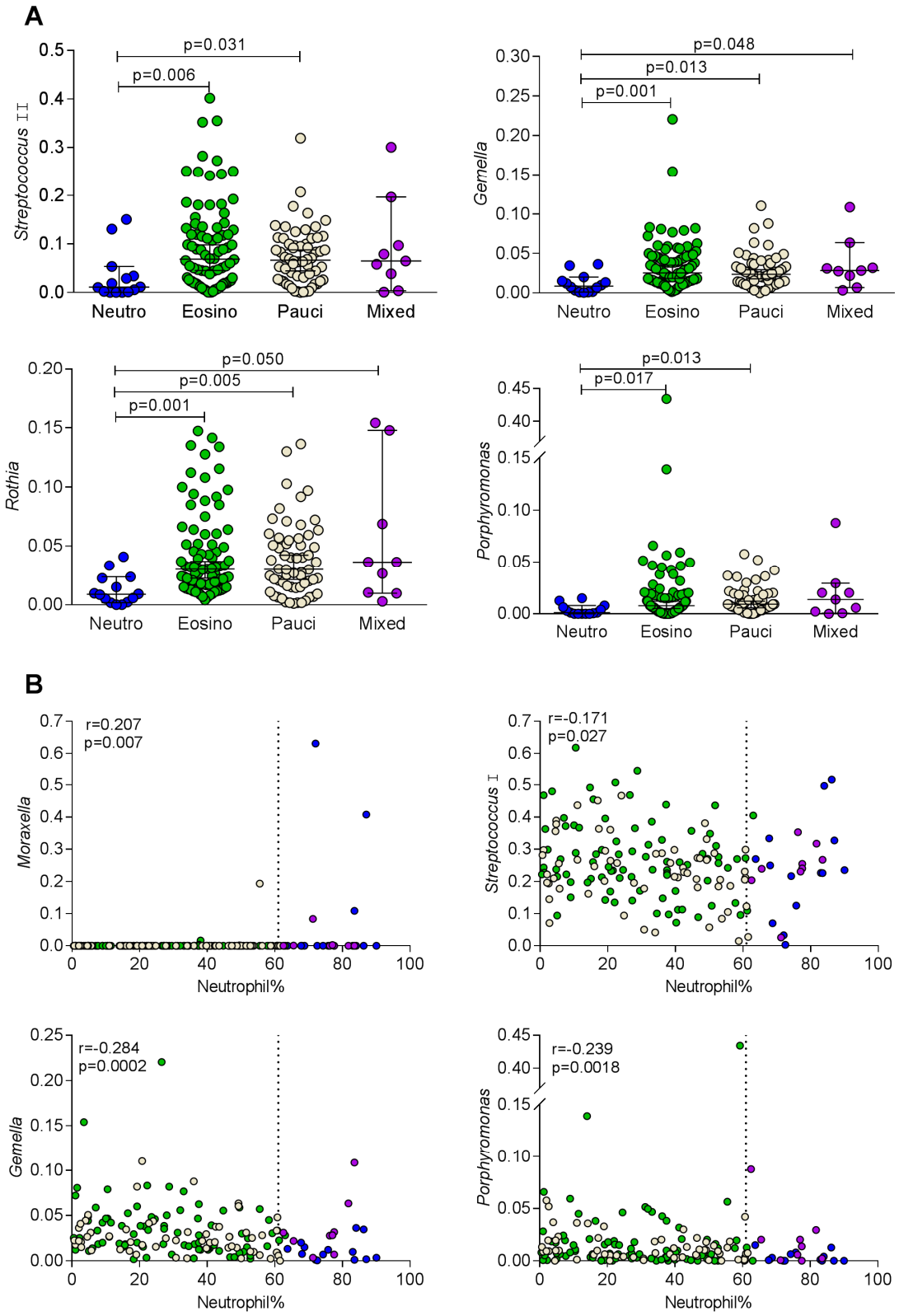
ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT


A
Streptococcus II
p=0.031
p=0.006
Neutro
Eosino
Pauci
Mixed
Gemella
p=0.048
p=0.013
p=0.001
Rothia
p=0.050
p=0.005
p=0.001
Porphyromonas
p=0.013
p=0.017
B
Moraxella
r=0.207
p=0.007
Neutrophil%
Streptococcus I
r=-0.171
p=0.027
Gemella
r=-0.284
p=0.0002
Porphyromonas
r=-0.239
p=0.0018

ACCEPTED MANUSCRIPT

Inflammatory phenotypes in severe asthma are associated with distinct airway microbiology

Steven L. Taylor*, Lex E.X. Leong*, Jocelyn M. Choo, Steve Wesselingh, Ian A. Yang, John W. Upham, Paul N. Reynolds, Sandra Hodge, Alan L. James, Christine Jenkins, Matthew J. Peters, Melissa Baraket, Guy B. Marks, Peter G. Gibson, Jodie L. Simpson†, Geraint B. Rogers†

Supplementary Methods

Exclusion criteria

Asthma diagnosis was established using American Thoracic Society guidelines based on current episodic respiratory symptoms, clinical diagnosis and evidence of variable airflow obstruction.[1] Participants with asthma were included if stable but symptomatic, despite being prescribed maintenance inhaled corticosteroid (ICS) and long acting bronchodilator treatment with an Asthma Control Questionnaire 6 (ACQ6) score >0.75.[2]

Participants with an FEV_1 <40% predicted, current smokers, ex-smokers who had ceased smoking in the previous year and those with a recent (past four weeks) exacerbation or respiratory infection were excluded. Those with significant smoking related air-space disease (ex-smokers >10 pack year history and DLCO/VA <70% predicted OR smoking history >10 pack years and exhaled carbon monoxide >10ppm) 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 centres

Sputum samples were collected from eight Australian centres:
Hunter Medical Research Institute, Newcastle NSW Australia
The Prince Charles Hospital, Chermside, QLD, Australia
Princess Alexandra Hospital, Woolloongabba QLD, Australia
Royal Adelaide Hospital, Adelaide SA, Australia
Sir Charles Gairdner Hospital, Nedlands WA, Australia
Woolcock Institute of Medical Research, Glebe NSW, Australia
Concord Repatriation General Hospital, Concord NSW, Australia
Liverpool Hospital, Liverpool NSW, Australia

Sample collection

All study participants attended a single visit that included the assessment of lung function asthma symptoms, asthma specific quality of life,[3] medication use and smoking status. Sputum induction with hypertonic saline (4.5%) was performed as described previously.[4] Sputum aliquots were stored at -80°C for DNA extraction, or dispersed using dithiothreitol for sputum cell count assessment and inflammatory subtype determination.[5]

Patient inflammatory phenotyping

Patient sputum was dispersed using dithiothreitol and inflammatory cells were counted as a percentage of total sputum cells. Inflammatory subtype was determined as described below. Neutrophilic cut-off values were age-dependent as described previously.[6, 7]

Neutrophilic phenotype

Neutrophil % (< 20 years old)	≥75.57%
Neutrophil % (20-40 years old)	≥61.61%

ACCEPTED MANUSCRIPT

Neutrophil % (40-60 years old)	≥63.25%
Neutrophil % (>60 years old)	≥67.25%
Eosinophilic phenotype	
Eosinophil %	≥3%
Paucigranulocytic phenotype	
Eosinophil %	≤3%
Neutrophil % (< 20 years old)	≤75.57%
Neutrophil % (20-40 years old)	≤61.61%
Neutrophil % (40-60 years old)	≤63.25%
Neutrophil % (>60 years old)	≤67.25%
Mixed granulocytic phenotype	
Eosinophil %	≥3%
Neutrophil % (< 20 years old)	≥75.57%
Neutrophil % (20-40 years old)	≥61.61%
Neutrophil % (40-60 years old)	≥63.25%
Neutrophil % (>60 years old)	≥67.25%

DNA extraction

DNA extraction was performed on sputum sample aliquots of approximately 100 µl. Following the addition of 300 µl of phosphate buffered saline, samples were vortexed for 10 seconds and placed on ice for 2 min. Bacterial cells were then pelleted by centrifugation at 13,000 x g for 10 min. Following removal of supernatant, 300 µl of Tris-EDTA solution (10 mM Tris-HCl, 1 mM EDTA; pH 8.0; Ambion, ThermoFisher Scientific, Victoria, Australia), 200 µg of silica: zirconium beads (1:1 of 0.1 µm and 1.0 µm; Biospec Products, Inc., OK, USA), and a single chrome bead (3.2 mm, Biospec Products, Inc., OK, USA) were added to the tube containing the cell pellet. Samples underwent bead-beating at 6.5 m/s for 60 sec in a FastPrep®-24 Instrument (MP Biomedicals, CA, USA). Homogenised sample was heated to 90 °C for 5 min, before being cooled on ice for 5 min. Lysozyme (ROCHE, ThermoFisher Scientific, Victoria, Australia) and lysostaphin (Sigma-Aldrich, MO, USA) were then added to a final concentration of 2 mg/mL and 0.1 mg/mL, respectively, and samples incubated at 37 °C for 1 hr. Proteinase K (Fermentas, ThermoFisher Scientific, Victoria, Australia) and sodium dodecyl sulphate (Sigma-Aldrich, MO, USA) were then added to a final concentration of 1.2 mg/mL and 1.5 %, w/v, respectively. Following incubation at 30 min at 56 °C, 40 µl of 5M sodium chloride and 450 µl of phenol:chloroform:isoamyl alcohol (25:24:1; saline buffered at pH8.0; Sigma-Aldrich, MO, USA) were added and samples vortexed for 30 sec. The aqueous-organic layers were separated by centrifugation at 13,000 x g for 10 min and 400 µl of the aqueous layer was transferred to a new microfuge tube. DNA was recovered using an EZ-10 Spin column in accordance with manufacturer's instructions (Bio Basic, Inc., Ontario, Canada), following precipitation by the addition of 10 M ammonium acetate and 99% ethanol (Sigma Aldrich, MO, USA) in a 1:10 and 1:1 ratio with sample volume, respectively. DNA was eluted in 50 µl UltraPure DNase/RNase-free distilled water (Gibco, ThermoFisher Scientific, Victoria, Australia) and stored at -80 °C prior to analysis.

16S rRNA gene amplicon sequencing

The V1-3 hypervariable region of the bacterial 16S rRNA gene was amplified from sputum DNA using modified primers 27F (5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGAGRGTTTGATCMTGGCTCAG-3') and 519R (5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTNTTACNGCGGCKGCTG-3'), with Illumina adapter overhang sequences as indicated by underline. Amplicons were generated, cleaned, indexed and sequenced according to the Illumina MiSeq 16S Metagenomic Sequencing Library Preparation protocol

ACCEPTED MANUSCRIPT
(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 µM), 5 µL of reverse primer (1 µM) and 12.5 µL of 2× KAPA HiFi Hotstart ReadyMix (KAPA Biosystems, Wilmington, MA, USA) in a total volume of 25 µL. The PCR reaction was performed on a Veriti 96-well Thermal Cycler (Life Technologies) using the following program: 95 °C for 3 min, followed by 25 cycles of 95 °C for 30 sec, 55 °C for 30 sec and 72 °C for 30 sec and a final extension step at 72 °C for 5 min. Samples were multiplexed using a dual-index approach with the Nextera XT Index kit (Illumina Inc., San Diego, CA, USA) according to the manufacturer's instructions. The final library was paired-end sequenced at 2 × 300 bp 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

Approximate 16S rRNA gene copy number was assessed by quantitative PCR (qPCR) using the 16S rRNA universal primers B331F (5'-TCCTACGGGAGGCAGCAGT-3') and B797R (5'-GGACTACCAGGGTATCTAATCCTGTT-3') using Platinum SYBR Green (ThermoFisher scientific, Vic, Australia) as previously described.[8] Reactions were performed in duplicate and averages taken. Sample total bacterial copy number was calculated per µL of DNA eluate against a standard curve of a known bacterial copy number.

Sequence data processing

The Quantitative Insights Into Microbial Ecology (QIIME, v1.8.0)[9] software was used to analyse the 16S rRNA sequence generated from paired-end amplicon sequencing using bioinformatics pipeline as previously described.[10] Briefly, barcoded forward and reverse sequencing reads were quality filtered and merged using Paired-End reAd mergeR (PEAR v0.9.6).[11] Chimeras were detected and filtered from the paired-end reads using USEARCH (v6.1)[12] against the 97% clustered representative sequences from the Greengenes database (v13.8).[13] Operational taxonomic units (OTUs) were assigned to the reads using an open reference approach with UCLUST algorithm (v1.2.22q) against the SILVA database release 111 (July 2012)[14] that was clustered at 97% identity. Spurious OTUs were then removed systematically using previous reports of common laboratory sequencing contaminants.[15] A minimum subsampling depth of 1,732 reads was then selected for all samples.

Where taxa assignment failed to classify to the Family or Genus level, OTU reference sequences (accounting for >99% of OTU reads) were separately aligned using SILVA Incremental Aligner (SINA) (https://www.arb-silva.de/) which uses SILVA, RDP, Greengenes, LTP and EMBL sequence collections. If the alignments identified taxa to a genus level, and at >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* during initial assignment.

Diversity measurements and statistical analyses

Bray-Curtis matrix was calculated based on sample-normalised, square root transformed relative taxon abundance. Principal coordinate analysis (PCoA) was used to visualize clustering of samples based on their similarity matrices with PCO1 and PCO2 coordinates and group centroids plotted using ggplot2 package of R statistical software.[16] Distance from centroid was calculated as previously described, using PRIMER.[17] Permutational multivariate analysis of variance (PERMANOVA)[18] on the beta-diversity matrices was used to test the null hypothesis of no difference amongst a priori-defined groups using PERMANOVA + add-on package for PRIMER. The test was computed using unrestricted permutation of raw data with 9,999 random permutations and at a significance level of 0.01.

Taxon dispersion

Variation in microbiota composition at the genus-level was assessed using multiple approaches. First, taxa that contributed to the overall variation between the asthma phenotypes were identified using SIMilarity of PERcentages (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 heatmap using ggplot2 package of R statistical software.[16] Hierarchical clustering of the taxa was performed on Bray-Curtis dissimilarity and clustered using single linkage method. Dominance of *Haemophilus* and *Moraxella* was determined when the relative abundance of each taxa exceeded 40%. This cut-off was selected based to the distribution of the relative abundance, where a clear distinction between samples with >40% and <40% was evident, suggesting overgrowth of these taxa.

Second, strong taxon-taxon correlations were identified using SparCC,[19] 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 ($p \leq 0.01$) using Cytoscape (version 3.4.0).

Two approaches were used to investigate the impact of pathogen overgrowth on microbiota composition. First, in samples where *Haemophilus* or *Moraxella* were the dominant taxon and represented $\geq$ 40% of total reads, their relative abundance was adjusted to the mean value for the study cohort and the remaining relative abundance measures rescaled, as described previously.[20] PERMANOVA analyses were then performed on the rescaled data. Second, pairwise comparisons between neutrophilic and eosinophilic samples were performed using the phyloseq R package[21] with DEseq2[22] extension, based on count data. p values were corrected using the Benjamini-Hochberg false discovery rate procedure and a corrected alpha value cut-off of <0.05 used for inclusion.

ACCEPTED MANUSCRIPT

References

1. Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. . *Am Rev Respir Dis* 1987; **136**(1): 225-44.
2. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. *Eur Resp J* 1999; **14**: 902-7.
3. Juniper EF, Guyatt GH, Ferrie PJ, Griffith LE. Measuring quality of life in asthma. *The American review of respiratory disease* 1993; **147**(4): 832-8.
4. 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. *American journal of respiratory and critical care medicine* 1998; **158**(1): 36-41.
5. Simpson JL, Phipps S, Baines KJ, Oreo KM, Gunawardhana L, Gibson PG. Elevated expression of the NLRP3 inflammasome in neutrophilic asthma. *The European respiratory journal* 2014; **43**(4): 1067-76.
6. Brooks CR, Gibson PG, Douwes J, Dalen CJV, Simpson JL. Relationship between airway neutrophilia and ageing in asthmatics and non-asthmatics. *Respirology* 2013; **18**(5): 857-65.
7. Simpson JL, Scott R, Boyle MJ, Gibson PG. Inflammatory subtypes in asthma: assessment and identification using induced sputum. *Respirology* 2006; **11**(1): 54-61.
8. 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**(1): 257-66.
9. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nature methods* 2010; **7**(5): 335-6.
10. 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.
11. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 2014; **30**(5): 614-20.
12. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010; **26**(19): 2460-1.
13. 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. *The ISME journal* 2012; **6**(3): 610-8.
14. 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 research* 2013; **41**(Database issue): D590-6.
15. 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 Biology* 2014; **12**(1): 1-12.
16. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, et al. gplots: Various R programming tools for plotting data. *R package version* 2009; **2**(4).
17. Anderson MJ, Ellingsen KE, McArdle BH. Multivariate dispersion as a measure of beta diversity. *Ecology Letters* 2006; **9**(6): 683-93.
18. Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 2001; **26**(1): 32-46.
19. Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. *PLoS computational biology* 2012; **8**(9): e1002687.
20. Rogers GB, van der Gast CJ, Serisier DJ. Predominant pathogen competition and core microbiota divergence in chronic airway infection. *The ISME journal* 2015; **9**(1): 217-25.
21. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one* 2013; **8**(4): e61217.
22. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 2014; **15**(12): 550.

ACCEPTED MANUSCRIPT

Supplementary Figure 1: Bacterial burden as assessed by 16S rRNA gene copy number. Bars show the median±95% CI. Statistical significance was assessed by Kruskal-Wallis one-way ANOVA with Dunn's post hoc test. No significant difference between phenotypes.

Supplementary Figure 2. Alpha diversity measures among asthma phenotypes. A) Taxa richness, B) Shannon-Wiener index, C) Simpson's evenness index, D) Pielou's evenness. Bars show the median±95% CI. Statistical significance was assessed by Kruskal-Wallis one-way ANOVA with Dunn's post hoc test.

Supplementary Figure 3: Correlations between sputum neutrophil/eosinophil counts (as a percentage of total cell count) and alpha diversity measures. A) Neutrophil % where dotted line at 61% neutrophils indicates phenotype cut-off point. B) Eosinophil % where dotted line at 3% eosinophils indicates phenotype cut-off point. Colours represent asthma phenotype, based on neutrophilia or eosinophilia where blue= >61% neutrophils, green= >3% eosinophils, yellow= <61% neutrophils and <3% eosinophils (paucigranulocytic), and purple= both >61% neutrophils and >3% eosinophils (mixed). Statistical significance was assessed by Spearman's rank correlation.

Supplementary Figure 4: Correlations between sputum neutrophil/eosinophil counts (as a percentage of total cell count) and weighted UniFrac distance from centroid. A) Neutrophil % where dotted line at 61% neutrophils indicates phenotype cut-off point. B) Eosinophil % where dotted line at 3% eosinophils indicates phenotype cut-off point. Colours represent asthma phenotype, based on neutrophilia or eosinophilia where blue= >61% neutrophils, green= >3% eosinophils, yellow= <61% neutrophils and <3% eosinophils (paucigranulocytic), and purple= both >61% neutrophils and >3% eosinophils (mixed). Statistical significance was assessed by Spearman's rank correlation.

ACCEPTED MANUSCRIPT

Supplementary Figure 5: A) Principal Coordinate Analysis (PCoA) of asthma phenotype groups based on Bray-Curtis similarity distances. The first two principal coordinates are plotted on the *x*- and *y*-axes, respectively (representing 36.5% of the total variation). B) Microbiota dispersion grouped by asthma phenotype. Distance from centroid calculated from Bray-Curtis dissimilarity matrix. C,D) Correlations between sputum inflammatory cell percentages and distance from centroid. C) Sputum neutrophil % vs Bray-Curtis distance from centroid. D) Sputum eosinophil % vs Bray-Curtis distance from centroid.

Supplementary Figure 6: Figure legend for bacterial network analysis. Showing weight and colour assigned to edges and nodes.

Supplementary Figure 7: Taxa which significantly correlated with eosinophil %. Colours represent asthma phenotype, based on neutrophilia or eosinophilia where blue= >61% neutrophils, green= >3% eosinophils, yellow= <61% neutrophils and <3% eosinophils (paucigranulocytic), and purple= both >61% neutrophils and >3% eosinophils (mixed). Dotted line at 3% eosinophils indicate phenotype cut-off points. Statistical significance was assessed by Spearman's rank correlation.

Supplementary Figure 8: Normalised, $\log_2$ fold changes of non-rarefied taxa read counts which significantly ($p<0.05$) differed between neutrophilic and eosinophilic phenotypes. Positive fold change indicates taxon which were significantly higher counts in neutrophilic participants, negative fold change indicates taxon which were significantly higher in eosinophilic participants. Showing that, when *Haemophilus* or *Moraxella* dominance do not influence data (due to non-rarefied count data as opposed to relative abundance), multiple taxa remain significantly different between neutrophilic and eosinophilic participants.

ACCEPTED MANUSCRIPT

Supplementary Table 1: 16S rRNA sequencing information

Median read count (Q1, Q3)	12792 (8060, 16595)
Subsample depth	1732
Samples excluded	7
Median Good's coverage (Q1, Q3)	0.952 (0.942, 0.963)

Supplementary Table 2: 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(perm)
Weighted UniFrac	Phenotype	3	9444.4	3148.1	3.9969	0.0004
	Residual	163	128390	787.64		
	Total	166	137830			
Bray-Curtis	Phenotype	3	8215.4	2738.5	3.3694	0.0001
	Residual	163	132480	812.74		
	Total	166	140690			

Supplementary Table 3: 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(perm)	Unique perms
Weighted UniFrac	Pauci vs Neutro	2.52	<0.01	9943
	Pauci vs Eosino	1.28	0.13	9928
	Pauci vs Mixed	0.80	0.62	9952
	Neutro vs Eosino	3.30	<0.0001	9924
	Neutro vs Mixed	1.20	0.21	9876
	Eosino vs Mixed	1.19	0.19	9939
Bray-Curtis	Neutro vs Eosino	2.89	<0.0001	9928
	Neutro vs Pauci	2.43	<0.001	9918
	Neutro vs Mixed	1.18	0.19	9888
	Eosino vs Pauci	1.15	0.17	9921
	Eosino vs Mixed	0.96	0.50	9899
	Pauci vs Mixed	0.78	0.83	9922

ACCEPTED MANUSCRIPT

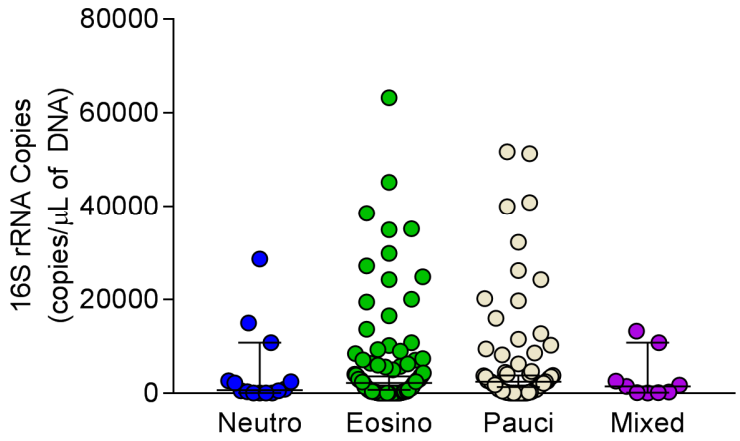
Supplementary Table 4: Similarity of percentages (SIMPER) analysis comparing taxa relative abundances between neutrophilic and eosinophilic phenotype groups. Showing thirteen top contributing taxa which collectively account for approximately 50% of variance between groups.

	Neutrophilic	Eosinophilic		
Species	Av. Abund	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 sp. oral clone DR002	0.05	0.08	1.1	2.17
Porphyromonas	0.04	0.11	1.1	2.17

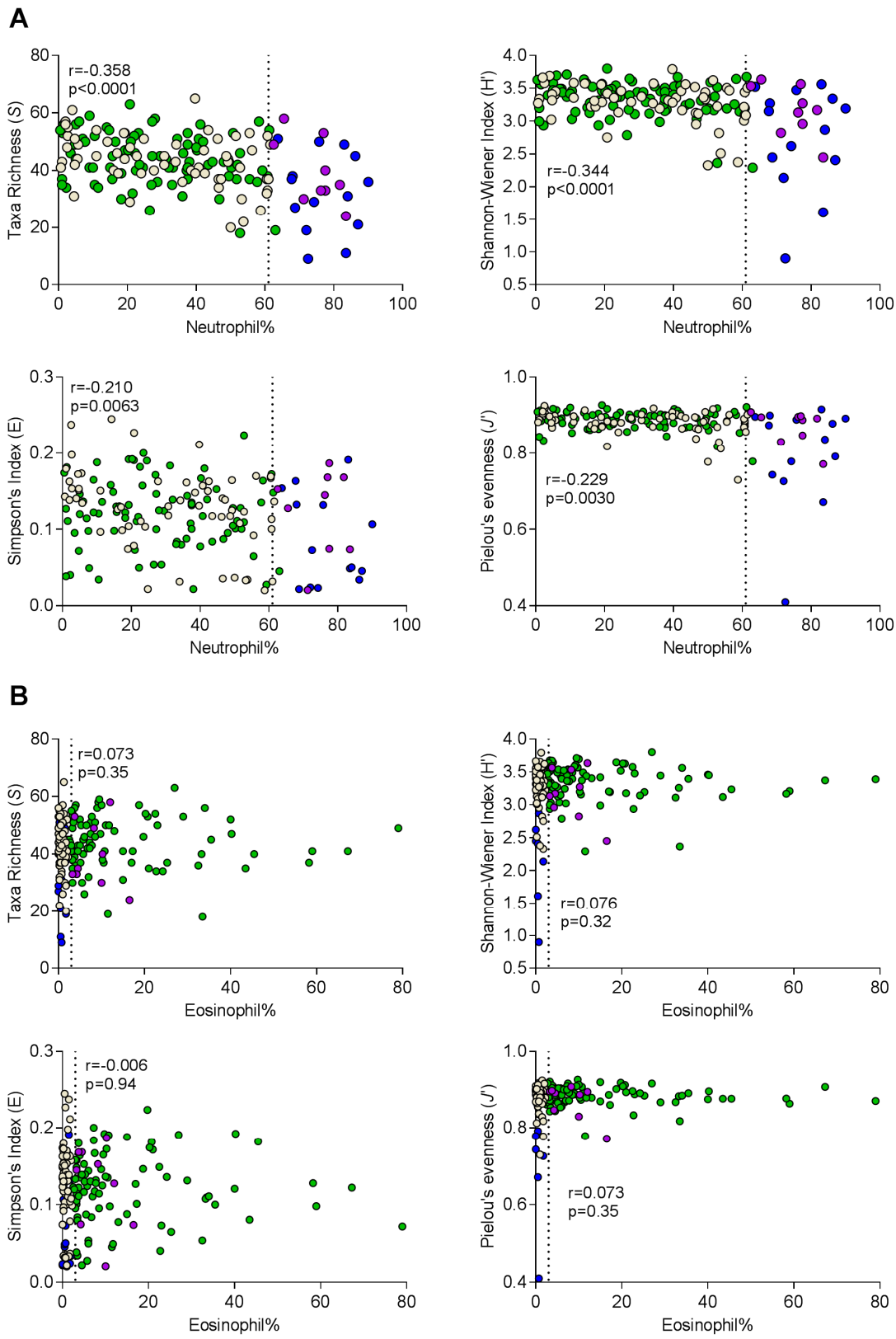
Supplementary Table 5: Top: PERMANOVA analysis, Bottom: Pairwise PERMANOVA on non-dominant microbiome. Bray-Curtis Distance of sputum microbiota on the genera level, grouped by asthma phenotype (permutations = 9999).

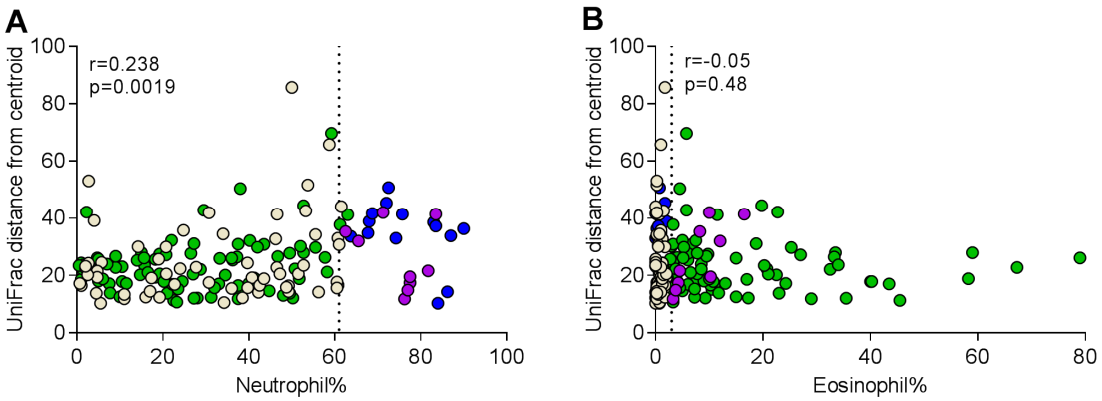
Source	**Df**	**SS**	**MS**	**Pseudo-F**	**P(perm)**
Phenotype	3	5381	1793.7	2.378	0.0004
Residual	163	122950	754.3		
Total	166	128330			

Groups	**T**	**P(perm)**	**Unique perms**
Neutro vs Eosino	2.31	<0.0001	9922
Neutro vs Pauci	2.26	<0.001	9911
Neutro vs Mixed	1.11	0.24	9857
Eosino vs Pauci	0.91	0.65	9904
Eosino vs Mixed	0.79	0.86	9921
Pauci vs Mixed	0.73	0.92	9924

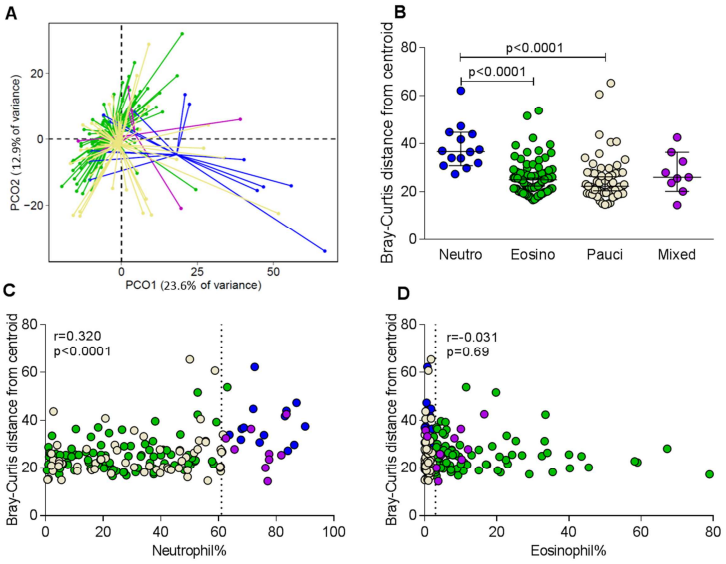


ACCEPTED MANUSCRIPT


A
Taxa Richness (S)
r=-0.358
p<0.0001
Neutrophil%
Shannon-Wiener Index (H')
r=-0.344
p<0.0001
Neutrophil%
Simpson's Index (E)
r=-0.210
p=0.0063
Neutrophil%
Pielou's evenness (J')
r=-0.229
p=0.0030
Neutrophil%
B
Taxa Richness (S)
r=0.073
p=0.35
Eosinophil%
Shannon-Wiener Index (H')
r=0.076
p=0.32
Eosinophil%
Simpson's Index (E)
r=-0.006
p=0.94
Eosinophil%
Pielou's evenness (J')
r=0.073
p=0.35
Eosinophil%



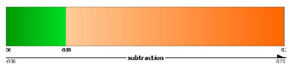
ACCEPTED MANUSCRIPT



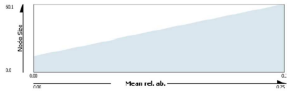
ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

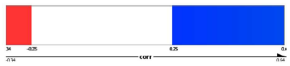
Node Fill Color Mapping



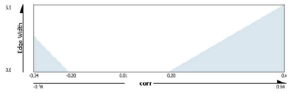
Node Size Mapping



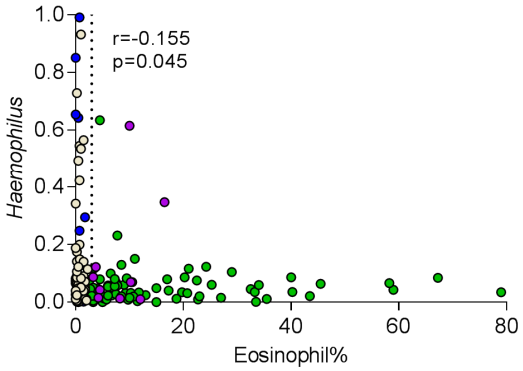
Edge Stroke Color (Unselected) Mapping

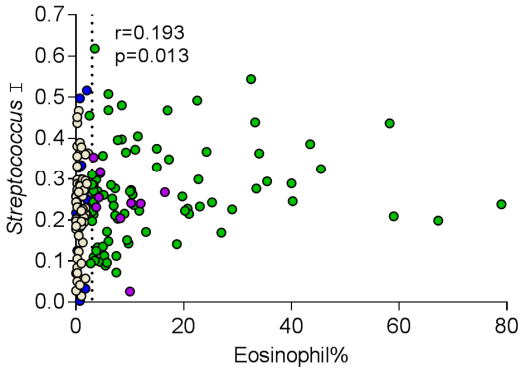


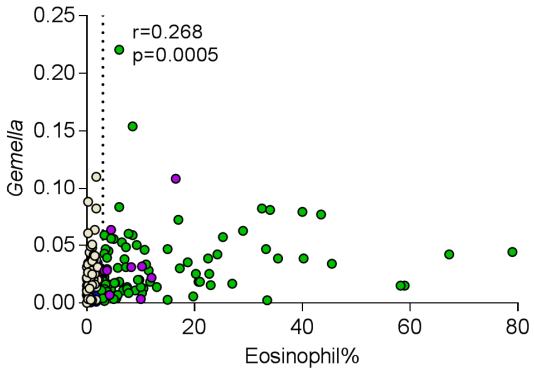
Edge Width Mapping

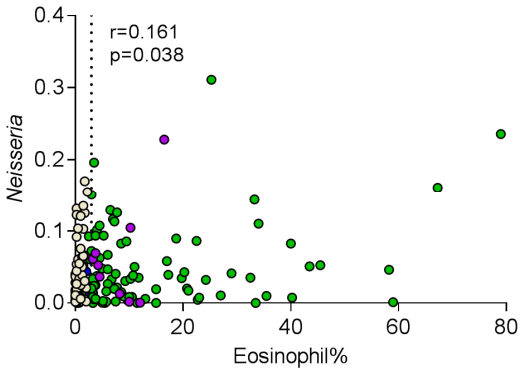


ACCEPTED MANUSCRIPT

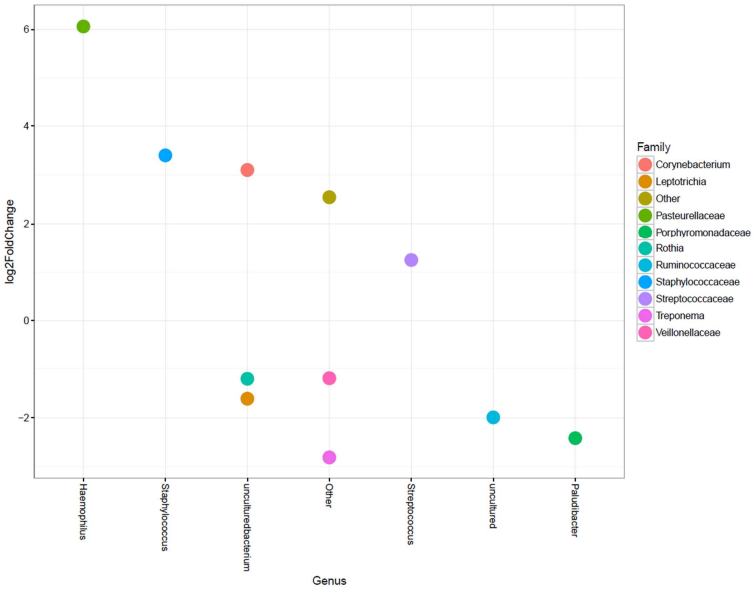

r=-0.155
p=0.045
Haemophilus
Eosinophil%


r=0.193
p=0.013
Streptococcus I
Eosinophil%

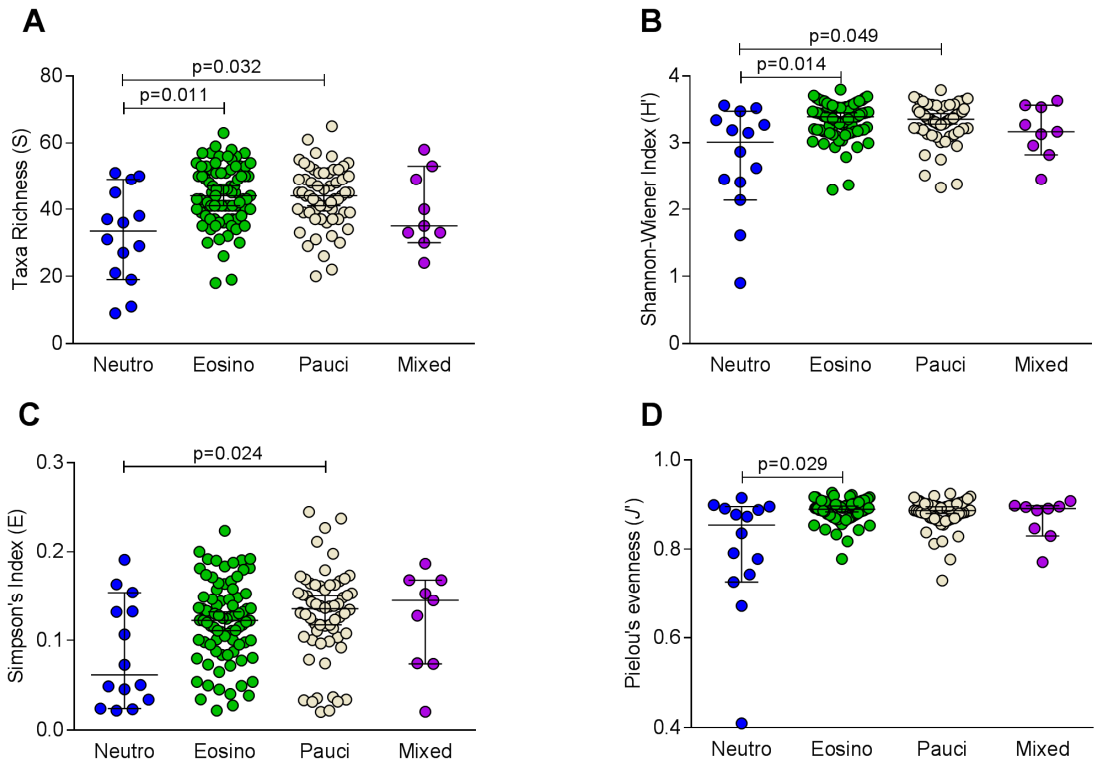

r=0.268
p=0.0005
Gemella
Eosinophil%


r=0.161
p=0.038
Neisseria
Eosinophil%

ACCEPTED



ACCEPTED MANUSCRIPT



ACCEPTED MANUSCRIPT