Sputum microbiome temporal variability and dysbiosis in chronic obstructive pulmonary disease exacerbations: an analysis of the COPDMAP study

Zhang Wang[1†], Richa Singh[2†], Bruce E. Miller[3], Ruth Tal-Singer[3], Stephanie Van Horn[4], Lynn Tomsho[4], Alexander Mackay[2], James P. Allinson[2], Adam J. Webb[5], Anthony J. Brookes[5], Leena M. George[6], Bethan Barker[6], Umme Kolsum[7], Louise E Donnelly[2], Kylie Belchamber[2], Peter J. Barnes[2], Dave Singh[7], Christopher E. Brightling[6], Gavin C. Donaldson[2], Jadwiga A. Wedzicha[2], James R. Brown[1*] on behalf of COPDMAP

[1] Computational Biology, Target Sciences, Research and Development (R&D), GlaxoSmithKline (GSK), Collegeville, PA 19426, USA
[2] National Heart and Lung Institute, Imperial College London, London, SW3 6NP, UK
[3] Respiratory Therapy Area Unit, R&D, GSK, King of Prussia, PA 19406, USA
[4] Target and Pathway Validation, Target Sciences, R&D, GSK, Collegeville, PA 19426, USA
[5] Department of Genetics, University of Leicester, Leicester, LE1 7RH, UK
[6] Institute for Lung Health, University of Leicester, Leicester, LE3 9QP, UK
[7] University of Manchester and University Hospital of South Manchester, Manchester, M23 9QZ, UK

[†] These authors contributed equally to the manuscript.

[*] Correspondence to:
James R. Brown (James.R.Brown@gsk.com)
1250 S. Collegeville Road, Collegeville, Pennsylvania, 19426-0989, United States
Mobile: +16102478580
Tel: +16109176374

Word count abstract: 241
Word count text body: 3283

Abstract

Background

Recent studies suggest that lung microbiome dysbiosis, the disease associated disruption of the lung microbial community, might play a key role in chronic obstructive pulmonary disease (COPD) exacerbations. However, characterizing temporal variability of the microbiome from large longitudinal COPD cohorts is needed to better understand this phenomenon.

Methods

We performed a 16S ribosomal RNA survey of microbiome on 716 sputum samples collected longitudinally at baseline and exacerbations from 281 COPD subjects at three UK clinical centres as part of the COPDMAP consortium.

Results

The microbiome composition was similar among centres and between stable and exacerbations except for a small significant decrease of *Veillonella* at exacerbations. The abundance of *Moraxella* was negatively associated with bacterial alpha diversity. Microbiomes were distinct between exacerbations associated with bacteria versus eosinophilic airway inflammation. Dysbiosis at exacerbations, measured as significant within subject deviation of microbial composition relative to baseline, was present in 41% of exacerbations. Dysbiosis was associated with increased exacerbation severity indicated by a greater fall in FEV1, FVC and a greater increase in CAT score, particularly in exacerbations with concurrent eosinophilic inflammation. There was a significant difference of temporal variability of microbial alpha and beta diversity among centres. The variation of beta diversity significantly decreased in those subjects with frequent historical exacerbations.

Conclusions

Microbial dysbiosis is a feature of some exacerbations and its presence, especially in concert with eosinophilic inflammation, is associated with more severe exacerbations indicated by a greater fall in lung function.

Key messages:

What is the key question?

How does the lung microbial community vary over time within COPD subjects and how is microbial dysbiosis in exacerbations implicated in disease characteristics?

What is the bottom line?

Dysbiosis of the sputum microbiome in COPD exacerbations, particularly in concert with eosinophilic inflammation, is associated with a greater decline in lung capacity during the exacerbation event.

Why read on?

The presented study entails the largest COPD sputum microbiome cohort to date with multiple study centres, aiming at in-depth examination of microbial temporal variability, dysbiosis, and disease phenotypes.

Introduction

Chronic obstructive pulmonary disease (COPD) is characterized by persistent symptoms and impaired lung function as a consequence of small airway obliteration and alveolar destruction, and is associated with chronic lung inflammation [1-3]. Acute exacerbations of COPD are a sudden onset of sustained worsening of these symptoms. Bacteria potentially play a key role in COPD pathogenesis [4 5], with respiratory bacterial pathogens such as *Haemophilus influenzae*, *Moraxella catarrhalis* and *Streptococcus pneumoniae* capable of driving host inflammatory responses [6-9]. Since bacteria frequently interact with each other and respond to altered environmental conditions, the consortium of the lung microbial community, known as the lung microbiome, could be important in the crosstalk between respiratory tract pathogens and host response [10 11].

Emerging studies collectively suggest that the lung microbiome differs between stable and exacerbations in COPD ([11-15], for review see [16]). For example, Molyneaux et al. found an increased representation of pathogenic *Proteobacteria* in particular *Haemophilus* in exacerbations following rhinovirus infection [12]. Huang et al. observed an increase of *Proteobacteria* during exacerbations with a predicted loss of function in maintenance of microbial homeostasis [13]. Recently, several of us published a longitudinal analysis of the sputum microbiome from 87 subjects from BEAT-COPD cohort [11]. Our analysis revealed an increased *Proteobacteria* versus *Firmicutes* during exacerbations. In addition, we found distinct microbiome composition between bacterial and eosinophilic exacerbations. In light of the heterogeneous nature of COPD exacerbations, the lung microbiome has potential as a biomarker to assist in the precision medicine treatments for specific COPD patient subpopulations.

Although insightful, results from these previous studies have limitations in terms of understanding microbial dysbiosis during exacerbations, as most of these studies comparing the microbiome at stable and exacerbations involved only one single sampling point of each state. The lung microbiome is temporally dynamic and can vary even in stable state [10].Thus the microbial changes during exacerbations are a mixture of both the disease associated disruption of microbial community or dysbiosis, and the regular temporal perturbations of the lung microbial composition. Therefore, examining the baseline variation of the lung microbiome is an important

first step to more precisely assess the extent of microbial dysbiosis during exacerbations. On the other hand, understanding temporal variability of the lung microbiome within individuals is also important in disease understanding. Disorder of the temporal balance of microbial ecosystem in the respiratory tract could trigger a dysregulated host immune response that results in negative effects on host biology [10]. Linking microbial temporal variation to disease characteristics and host inflammatory profiles could potentially lead to monitoring and manipulating the stability of airway microbial composition as a therapeutic strategy for COPD.

A finer-grained longitudinal sampling of microbiome at multiple stable and exacerbation visits is necessary to quantitatively measure temporal variability of the microbiome and assess the significance of microbial dysbiosis during exacerbations. Here we describe a longitudinal 16S ribosomal RNA (rRNA) gene based microbiome survey on 716 sputum samples collected sequentially at baseline and exacerbations over a period of up to two years duration from 281 COPD subjects at three UK centres as part of the COPDMAP consortium. This entails one of the largest COPD sputum microbiome cohorts to date aiming at in-depth examination of temporal variability of the microbiome. We provide new insights into temporal changes of the microbiome and its potential implication in disease progression.

Material and Methods

Subjects and samples

Full information on subject inclusion/exclusion criteria, sputum sample collection, microbiome and statistical analyses are provided in the online supplementary appendix. Briefly, sputum samples were collected at multiple longitudinal baseline and exacerbation visits from COPD subjects at three clinical centres, Imperial College London, University of Leicester and University Hospital of South Manchester (hereafter referred to as London, Leicester and Manchester, respectively) as part of the COPDMAP consortium (www.copdmap.org). All sputum samples were immediately stored at -80ºC and shipped frozen in batches for analysis. Exacerbations were treated with corticosteroids and antibiotics according to guidelines [17]. The protocol summary is available at https://clinicaltrials.gov/ (Identifier: NCT01620645).

Microbiome analysis

For quality control purposes, all DNA extractions, sequencing and data analyses were performed in a single, centralized lab at the GSK R&D facility in Collegeville, Pennsylvania, USA. Bacterial genomic DNA was extracted from frozen sputum samples using the Qiagen DNA Mini kit (Qiagen, CA, USA) as per manufacture protocol. The V4 hypervariable region of the 16S rRNA gene was PCR amplified and sequenced using multiplexed Illumina Miseq platform with the proper controls against reagent contamination as described previously [11]. Sequencing reads were processed using QIIME pipeline version 1.9 [18]. The default set of criteria was used to remove low quality and chimeric reads. The remaining reads were subject to a close reference OTU picking (97% identity cutoff). Sequence data are deposited at the National Centre for Biotechnology Information Sequence Read Archive (SRP102480).

Statistical analysis

Exacerbation phenotypes were defined using microbiological and clinical criteria as established previously [12]. Phenotypes of 146 exacerbations samples were undetermined due to missing data. Partial Least Squares Discriminant Analysis (PLS-DA) was performed on exacerbation phenotypes and microbiome and/or clinical data using SIMCA-P (Umetrics, Stockholm, Sweden) [19]. Dysbiosis at exacerbations was measured as the deviation (Z-score) of the first

153 Principal Coordinate (PC1) of the weighted UniFrac distance for exacerbation samples relative to
154 all baseline PC1s from the same subject. Temporal variability of microbial alpha and beta
155 diversity was measured using the metrics described by Flores et al. [20]. A general linear model
156 (GLM) was constructed between demographic and baseline clinical variables and temporal
157 variability of alpha and beta diversity among subjects. The model was optimized in a stepwise
158 algorithm using the "step" function in the R stats package [21]. The false discovery rate (FDR)
159 method was used to adjust *P*-values for multiple testing wherever applicable [22].
160

Results

Overview of the COPDMAP sputum microbiome

Microbial composition was determined for 716 sputum samples collected at baseline and exacerbations from 281 COPDMAP subjects at three centres. The number of samples varies from one to nine per subject (Fig. S1). Demographic and baseline clinical data were recorded for subjects at initiation of sample collection (Table 1, Table S1). A set of 16 clinical and biochemical characteristics were further collected longitudinally (Table 2, Table S2). From DNA sequences of the V4 hypervariable region of the 16S rRNA gene, a total of 3,784 operational taxonomic units (OTUs) were identified using 97% identity cut-off after rarefaction.

Table 1. Major demographic and baseline clinical features of all subjects.

Demographic and baseline features	All subjects (N=281) *
Gender †	Male: 187 (70.3%), Female: 79 (29.7%)
Age ‡	70 (8.1)
BMI	27.2 (5.4)
Baseline GOLD status	1: 30 (11.4%), 2: 132 (50.2%), 3: 78 (29.7%), 4: 23 (8.7%)
Treatment #	Antibiotics: 38 (15.3%), Steroids: 9 (3.6%), Both: 202 (81.1%)
Number of cigarette packs per year [1]	47 (30)
Number of exacerbation per year [1]	1.1 (1.6)
Baseline FEV1	1.5 (0.6)
Baseline FEV1%	56.3 (18.9)
Baseline FEV1 predicted	2.6 (0.5)
Baseline FVC	2.9 (1.0)
Baseline FEV1/FVC ratio	0.5 (0.1)
CAT score	18.7 (7.3)
CES-D score [1]	10 (13)

SGRQ total score	47.4 (18.2)

† Categorical data present as number (proportion). ‡ Continuous data present as mean (SD) unless stated below.

[1] Median (interquartile range).

* 15 subjects were missing any demographic or clinical data.

[#] The numbers represent exacerbation events, thus include subjects with more than one exacerbation.

Table 2. Major longitudinal clinical features at baseline and exacerbations of all samples.

Longitudinal features	All samples (N=716)	Visits		*P*-value ‡
		Baseline (N=446)	Exacerbations (N=270)	
FEV1 †	1.4 (0.5)	1.5 (0.5)	1.2 (0.5)	<0.001
FVC	2.8 (0.9)	3.0 (0.8)	2.5 (0.9)	<0.001
FEV1/FVC ratio	0.5 (0.2)	0.5 (0.2)	0.5 (0.2)	0.26
CAT score	21.1 (7.4)	19.6 (7.1)	24.2 (7.0)	<0.001
C-reactive protein (CRP) [1]	5.0 (11.0)	3.0 (5.0)	10.0 (27.0)	<0.001 [2]
Blood neutrophil count ($X10^9$ cells/L)	5.5 (2.3)	4.9 (1.7)	6.2 (2.7)	<0.001
Blood lymphocyte count ($X10^9$ cells/L)	1.8 (0.7)	1.8 (0.6)	1.8 (0.7)	0.49
Blood monocyte count ($X10^9$ cells/L)	0.7 (0.3)	0.6 (0.2)	0.7 (0.3)	<0.001
Blood eosinophil count ($X10^9$ cells/L) [1]	0.2 (0.2)	0.2 (0.2)	0.2 (0.2)	0.18 [2]
Blood basophil count ($X10^9$ cells/L)	0.0 (0.0)	0.1 (0.0)	0.0 (0.0)	0.01
Sputum neutrophil count % [1]	78.8 (33.8)	75.1 (34.0)	84.2 (28.5)	<0.001 [2]
Sputum lymphocyte count % [1]	0.0 (0.5)	0.0 (0.3)	0.2 (1.0)	0.028 [2]
Sputum eosinophil count % [1]	0.8 (2.0)	0.8 (2.2)	0.5 (2.0)	0.07 [2]
Sputum macrophage count % [1]	13.0 (21.2)	14.5 (23.2)	8.5 (19.0)	<0.001 [2]
Sputum epithelial count % [1]	3.2 (8.0)	4.0 (9.8)	2.0 (4.8)	<0.001 [2]

† Data present as mean (SD) unless stated below.

[1] Median (interquartile range).

‡ P-value was calculated for baseline and exacerbations comparison using T-test unless stated below.

[2] Mann-Whitney-Wilcoxon Test.

Similar to other sputum or lung microbiome studies [11-15 23-26], the vast majority of OTUs belonged to *Proteobacteria* (52.3%), *Firmicutes* (28.7%), *Bacteroidetes* (15.0%) and *Actinobacteria* (1.9%) at the phylum level (Table S3, Fig. S2). At the genus level, *Haemophilus* (25.8%) was most abundant across all samples, followed by *Veillonella* (15.8%) and *Prevotella* (13.2%). Other common genera in the airway such as *Streptococcus* (4.4%) and *Moraxella* (4.0%) were also among the most abundant genera identified. As a quality control for sample processing and sequence analyses, an additional aliquot of sputum was collected as duplicates for 11 samples from the same subject at the same visit. Duplicates all had low UniFrac distance and were highly similar in microbial composition (Fig. S3).

Overall, the microbiome composition was similar between baseline and exacerbation samples with a small significant decrease of *Veillonella* at exacerbations (repeated measures ANOVA, FDR-adjusted (adj.) $P = 0.042$) (Fig. 1A). The microbiome composition was similar among centres with a significantly higher alpha diversity in the London cohort (Fig. S4A). Within each centre, there was a significant decrease of alpha diversity (Shannon, repeated measures ANOVA, P=1.1e-4) and a non-significant increase of *Moraxella* (repeated measures ANOVA, adj. P= 0.092) at Leicester (Fig. S4B). A strong negative correlation was found between the abundance of *Moraxella* and alpha diversity for all samples (Shannon, R=-0.445, adj. P<9.6e-14, Fig. 1B).

Similar to previously observed [11], distinct microbial populations were found in bacterial and eosinophilic exacerbations, with a significantly decreased alpha diversity (Shannon, T-test P=0.008) and significantly increased proportion of *Proteobacteria* (T-test, adj. P=0.001) versus *Bacteroidetes* (T-test, adj. P=0.002) in bacterial exacerbations compared to eosinophilic exacerbations (Fig. 1C-D, Fig. S5A). An improvement in predicting the two phenotypes was observed according to PLS-DA by combining the clinical and microbiome datasets versus using the clinical data only (Fig. S5B). Within individual centres, this trend was more pronounced for Leicester samples than those of London or Manchester (Fig. S5C).

We performed multivariate analysis to identify clinical factors significantly associated with microbial alpha and beta diversity. Among all clinical variables, C-reactive protein (CRP), a

known inflammatory marker for COPD prognosis [27], was the most significant factor correlated with both alpha diversity (Shannon, $P<0.01$, Fig. S6) and beta diversity at the phylum level among all samples (Table S4). No factors significantly predicted variation at the genus and OTU levels. CRP was also significantly associated with alpha and beta diversity of the predicted functional profiles of the sputum microbiome using the software PICRUSt [28] (Table S5).

Increased disease severity in exacerbations with dysbiosis

To explore variation of the sputum microbiome over time, we plotted the first Principal Coordinate (PC1) of the weighted UniFrac distance for all samples within each subject as a proxy for their microbial compositions, as it explains 49.0% of the total beta diversity (Fig. 2A). Only subjects with at least two baseline and one exacerbation samples were included. Visual inspection of the plot revealed a deviation of PC1 for many exacerbation samples relative to baseline samples from the same subject (Fig. 2A, Fig. S7A), indicating specific exacerbations were particularly susceptible to alternation of microbial composition or dysbiosis. In comparison, the sputum microbiome was much less variable among baseline samples. This is supported by a significantly increased within subject standard deviation of PC1 (paired T-test, P=6.7e-4) combining baseline and exacerbation samples compared to baseline samples only, with the most profound changes at the Leicester centre (Fig. S7B).

Having assessed temporal variability of the sputum microbiome at baseline, we measured the dysbiosis of exacerbation as a Z-score that measures how much its PC1 deviated from all baseline PC1s from the same individual. A total of 49 exacerbations (out of 119 exacerbations with a Z-score, 41.2%) were identified as in significant dysbiosis state with an absolute Z-score greater than 2 ($P<0.05$, Fig. 2A, Fig. S7C). In most of these exacerbations, the sputum microbiome shifted from a balanced composition to a more biased one predominated by one or a few taxa with a decreased alpha diversity (Fig. S8). Bacterial genera of *Veillonella*, *Cronobacter* and *Haemophilus* were among the key taxa associated with the dysbiosis (Fig. S9). Across all exacerbation subtypes, bacterial exacerbations had the highest number of dysbiosis events than other subtypes (Fig. S7C), with the caveat that phenotype could not be defined for 21 of the 49 exacerbations due to missing data.

For exacerbations with or without significant dysbiosis, we compared the exacerbation severity determined by change in lung function and symptoms relative to the last baseline measurement. We found a non-significantly greater decrease in FEV1 and FVC and a greater increase in CAT score for exacerbations with dysbiosis compared to those without (Fig. 2B). Such trends were overall consistent within each centre, except for a reversal trend of FVC in Manchester which has a smaller sample size (Fig. S10A). Also, the exacerbation Z-score was positively correlated, albeit non-significantly, with changes of FEV1 and FVC, and negatively correlated with change of CAT score (Fig. S10B), suggesting that the more dysbiotic the exacerbation was, the more severe the clinical outcome could possibly be. As eosinophil abundance is another important factor for COPD exacerbations, we reclassified exacerbations according to both the dysbiosis and blood eosinophil indices. Doing so revealed four subgroups of exacerbations where dysbiosis and/or high blood eosinophil level (>3 $x10^8$ cells/L) are the predominant feature. Exacerbations with both dysbiosis and high eosinophil level had the greatest changes of FEV1 (statistically significant, ANOVA P=0.02), FVC and CAT score, whereas exacerbations with neither dysbiosis nor high eosinophils level were associated with the least of such changes (Fig. 2B).

Exacerbation frequency associated with temporal variability of the sputum microbiome

We next sought to quantify temporal variability of the sputum microbiome within subjects using the metrics described by Flores et al. [20]. Only subjects with at least three samples were included. The variability of microbial alpha diversity was denoted as the coefficient of variation of Shannon for samples within each subject. The variability of beta diversity was calculated as the median of pairwise UniFrac distances for samples within each subject. A wide range of temporal variability of alpha and beta diversity was observed across subjects (Fig. 3A). We noted that there was a significantly lower variation of alpha and beta diversity among London subjects than Leicester or Manchester ones (T-test, P<0.001). As expected, both variations of alpha and beta diversity were significantly higher in subjects with dysbiosis exacerbations than those without (T-test, P<0.01).

We constructed a generalized linear model (GLM) to look for clinical characteristics associated with temporal variability of the sputum microbiome. A set of 14 demographic and baseline clinical variables were included for each subject. Centre, FEV1/FVC ratio and number of exacerbations per year (prior to the sampling visits) were significant factors for the variation of alpha diversity across all subjects (Table 3). When reconstructing GLM for each centre, number of exacerbations per year was significant for London and Leicester subjects. In addition, centre and historical number of exacerbations per year were significantly associated with beta diversity variation across all subjects (Table 2). A continuous decreasing trend of historical number of exacerbations per year was observed toward subjects with greater variation of beta diversity (Fig. 3B). Likewise, a continuous decreasing trend of beta diversity variation was observed toward subjects with higher historical number of exacerbations per year (ANOVA, $P<0.05$, Fig. 3C). Similar trends were observed within each centre (Fig. S11) and for temporal variability of baseline microbiomes only (Fig. S12), although the association of baseline microbiome variability was not statistically significant (variability of beta diversity: $P=0.088$, variability of alpha diversity: $P=0.249$).

Table 3. List of demographic and baseline clinical variables significantly associated with temporal variability of microbial alpha and beta diversity among subjects. P-values are indicated for variables in the model. Significant variables are highlighted in asterisks.

Temporal variability	**Alpha diversity (Shannon)**				**Beta diversity (Weighted UniFrac distance)**			
	All	**London**	**Leicester**	**Manchester**	**All**	**London**	**Leicester**	**Manchester**
Historical number of exacerbations per year	0.01*	0.01*	3E-4*	0.32	0.01*	0.03*	0.47	0.21
BMI	0.11 ‡	0.73	0.01*	0.53	0.51	0.70	0.04*	0.88
CES-D score	0.50	0.03	0.35	0.68	0.44	0.32	0.63	0.01*
Packs of cigarette per year	0.40	0.67	0.31	0.38	0.07‡	0.50	0.02*	0.58
FEV1	0.03	0.76	0.59	0.82	0.64	0.95	0.17‡	0.41
FEV1/FVC ratio	0.65	0.88	0.01*	0.40	0.87	0.36	0.77	0.98
Age	0.54	0.03*	0.38	0.23	0.70	0.37	0.99	0.26
SGRQ total score	0.61	0.92	0.80	0.38	0.33	0.44	0.40	0.03*
Centre	2.6E-12*	NA	NA	NA	1.8E-10*	NA	NA	NA

‡ Variables not statistically significant but present in the model.

Discussion

Culture-independent analyses have uncovered a previously unappreciated complexity of the lung bacterial community that has reshaped our understanding of COPD aetiology [11-15 23-26]. Our study reveals a diverse sputum microbiome among the COPDMAP subjects and further validates the association of microbiome with specific exacerbation phenotypes. We also show in-depth temporal variation of the sputum microbial community within subjects and identified potential new relationships of the microbiome variation with patient disease progression.

One advantage of our study is the longitudinal sampling at multiple baseline and exacerbation visits compared to most previous studies where a single snapshot of exacerbations was taken. It is well appreciated that the lung microbiome is inherently variable shaped by the balance of ecological factors like microbial immigration and elimination [10]. During exacerbations, the balance goes awry with dysregulated host immune response and inflammation leading to further microbial changes or dysbiosis. Therefore, to explicitly determine the extent of disease associated dysbiosis one would need to first distil the normal perturbations of microbial composition. Our study underscores the importance of considering temporal variability of the microbiome in understanding the significance of microbial dysbiosis in COPD exacerbations.

From assessing temporal variation of the sputum microbiome, we identified a subset of exacerbation events in which significant dysbiosis is a feature. In these exacerbations, the microbiome composition shifted from a highly diverse microbial community to a less diverse one characterized by the predominance of only one or few genera. These dysbiosis exacerbations appear to be the main source of microbial temporal variation and are associated with a greater worsening of health status and decrease of lung capacity. To our knowledge, this is the first evidence to suggest that respiratory dysbiosis is associated with increased exacerbation severity in COPD, although the strength of this association is weak and needs to be further validated in additional cohorts and by other measures of disease severity. Altered environmental conditions in exacerbations could disturb the composition of the lung microbial community [29 30], which in turn elicit a dysregulated host immune response through bacterial metabolites and virulent factors, resulting in a sustained damage cycle with an accelerated decline in lung function [31].

Whether dysbiosis is the cause or consequence of the increased exacerbation severity and how this imbalance is implicated in host inflammatory pathways are new questions that will impact on how we understand and treat COPD exacerbations.

It has been recently emphasized that not all COPD exacerbations are the same [32]. Our results suggest the existence of subgroups of exacerbations associated with or without significant microbial dysbiosis or increased eosinophilia. Importantly, these subgroups likely reflect fundamental differences in their immuno-pathogenesis driving the exacerbations, and therefore might require alternative therapeutic approaches. Interestingly, the most severe exacerbations were observed in the small subgroup that had evidence of bacterial dysbiosis in concert with eosinophilic inflammation. It is possible that this group might require interventions such as antibiotics and steroids (i.e. prednisolone) to target both bacteria and eosinophilic inflammation whereas in contrast those without bacterial dysbiosis nor eosinophilic inflammation might not require these therapies. Our results perhaps establish a new paradigm in stratifying COPD exacerbations according to dysbiosis and eosinophil measurements, which could be informative guiding future personalized therapies. Further efforts in identifying biomarkers for these subgroups in larger populations could help refine exacerbation subtypes toward phenotype-specific clinical management.

We found a significantly decreased historical exacerbation frequency in subjects with higher temporal variation of the microbial beta diversity. In COPD there is a subset of frequent exacerbators that are particularly susceptible to recurrent exacerbations independent of other risk factors [33 34]. Thus, low temporal variability of the sputum microbiome might come as a predicative factor for the frequent exacerbator phenotypes. We observed that a rise in exacerbation frequency was associated with a decline in the variation of microbiome beta diversity. This trend was statistically significant in the analysis of all microbiome samples but non-significant for that of a subset comprised of only baseline samples. Therefore, the predictability of sputum microbiome variability for exacerbation frequency still warrants further validation in larger longitudinal cohort studies.

An important novelty of our study is that there were three unique study centres. All samples were processed in a central lab, which minimizes microbiome variation due to differences in experimental protocols. Interestingly, there was a significant difference in temporal variability of the microbiome among subjects in the three centres, even though their overall microbiome profiles were highly similar. Factors accounting for the among-centre variation could include differences in the frequency of clinical visits and compliance with medications, although we lack the comparative data across centres to suggest specific causes. Our study suggests that the impact of demographics and clinical procedures on the lung microbial community needs to be broadly considered in future studies.

Our study has several caveats. First, only a proportion of the bacterial 16S rRNA gene was sequenced to characterize the microbial population both here and in previous lung microbiome studies [11-14]. Thus the resolution is insufficient when it comes to species-level characterization of the microbiome, whereas ecological and functional interaction of individual species or strains could be important in the underlying disease aetiology. Second, despite a large cohort size, longitudinal sampling remains relatively sparse for many subjects with variation in the timing of their sampling visits. Further efforts on characterizing respiratory tract metagenomes in a more regularly and intensively followed patient cohort together with host multi-omics profiling would promise to bring in a more comprehensive picture of the intrinsic variability of the lung microbiome and its implications in disease heterogeneity.

In summary, our study revealed a temporally dynamic sputum microbiome in COPD subjects in which microbial dysbiosis in exacerbations, particularly in concert with eosinophilic inflammation, was associated with increased exacerbation severity. Our findings underscore the importance of considering temporal variability of the sputum microbiome in COPD heterogeneity and its potential as a biomarker toward more precise treatment of COPD.

Acknowledgments

The study was funded by the Medical Research Council (UK) and COPD MAP (MRC/ABPI Inflammation and Immunology Initiative). The work was supported by the Leicester NIHR

BRC, University of Leicester, Leicester UK. This article presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Contributions

ZW, RS, BEM, RT-S, LED, KB, PJB, DS, CEB, GCD, JAW and JRB contributed to the study conception and design. AM, JPA, AJW, AJB, LMG, BB, UK, LED, KB, DS, CEB, GCD and JAW coordinated the collection of sputum samples and clinical data. SVH and LT performed microbiome DNA purification and sequencing. ZW performed microbiome computational analyses. AJB performed other statistical analysis. ZW wrote the initial draft of the manuscript with additional content provided by RS, RT-S, PJB, CEB and JRB and critical revisions from all authors. All authors read and approved the final version of the manuscript.

Figure legends

Figure 1. Baseline and exacerbation microbiome profiles across centres. A) Alpha diversity (Shannon) and compositions of major phyla and genera in samples at baseline and exacerbations. B) Correlation between alpha diversity (Shannon) and relative abundance of *Moraxella*. Each dot represents a sample coloured by baseline or exacerbations. C) Alpha diversity (Shannon) and composition of major phyla and genera in exacerbation samples of different exacerbation phenotypes. D) Principal Coordinate Analysis (PCoA) showing distinct clustering of samples with bacterial and eosinophilic exacerbations. The number of samples is indicated in the parenthesis under each subgroup in the bar chart. B: bacterial; V: viral; E: eosinophilic; BE: bacterial and eosinophilic; BV: bacterial and viral; and Pauci: pauci-inflammatory. Error bars are within 1.5 interquartile range of the upper and lower quartiles. *** adj. $P<0.001$; ** adj. $P<0.01$; * adj. $P<0.05$.

Figure 2. Dysbiosis of the sputum microbiome. A) Scatter plot of the first Principal Coordinate (PC1) of all samples within each subject at each centre. Only subjects with at least two baseline and one exacerbation samples were included. Exacerbation samples are highlighted in red. Box-whisker plots indicate the distribution of baseline PC1s within each subject. The confidence bands indicate the 95% confidence interval for the mean baseline PC1s within each subject. Exacerbations outside the confidence bands are the ones with significant dysbiosis (absolute Z-score>2, $P<0.05$). B) Box-whisker plots showing changes of FEV1, FVC and CAT score between dysbiosis and non-dysbiosis exacerbations, and among four subgroups of exacerbations classified by dysbiosis and blood eosinophils level. Error bars are within 1.5 interquartile range of the upper and lower quartiles.

Figure 3. Temporal variability of the sputum microbiome. A) Temporal variability of microbial alpha (coefficient of variation of Shannon) and beta diversity (median of pairwise weighted UniFrac distances) for each subject. Only subjects with at least three samples were included. Subjects with dysbiosis exacerbations are highlighted in yellow. B) Box-whisker plots showing exacerbation frequency of subjects within different quartile groups of temporal variability of alpha and beta diversity, with the first quartile defined as ‘low’, the second and third quartiles as

426 ‘medium’ and the fourth quartile as ‘high’. C) Box-whisker plots showing temporal variability of
427 alpha and beta diversity in subjects with different classes of exacerbation frequency. The number
428 of samples is indicated in the parenthesis under each subgroup in the box-whisker plot. Error
429 bars are within 1.5 interquartile range of the upper and lower quartiles. ANOVA test for
430 temporal variability of alpha and beta diversity: *** adj. $P<0.001$; ** adj. $P<0.01$; * adj. $P<0.05$.
431

Reference

1. Lopez AD, Shibuya K, Rao C, et al. Chronic obstructive pulmonary disease: current burden and future projections. *The European respiratory journal* 2006;27(2):397-412. doi: 10.1183/09031936.06.00025805
2. Franklin W, Lowell FC, Michelson AL, et al. Chronic obstructive pulmonary emphysema; a disease of smokers. *Annals of internal medicine* 1956;45(2):268-74.
3. Taraseviciene-Stewart L, Douglas IS, Nana-Sinkam PS, et al. Is alveolar destruction and emphysema in chronic obstructive pulmonary disease an immune disease? *Proceedings of the American Thoracic Society* 2006;3(8):687-90. doi: 10.1513/pats.200605-105SF
4. Erkan L, Uzun O, Findik S, et al. Role of bacteria in acute exacerbations of chronic obstructive pulmonary disease. *International journal of chronic obstructive pulmonary disease* 2008;3(3):463-7.
5. Sethi S, Murphy TF. Bacterial infection in chronic obstructive pulmonary disease in 2000: a state-of-the-art review. *Clinical microbiology reviews* 2001;14(2):336-63. doi: 10.1128/CMR.14.2.336-363.2001
6. Miravitlles M, Espinosa C, Fernandez-Laso E, et al. Relationship between bacterial flora in sputum and functional impairment in patients with acute exacerbations of COPD. Study Group of Bacterial Infection in COPD. *Chest* 1999;116(1):40-6.
7. Ball P. Epidemiology and treatment of chronic bronchitis and its exacerbations. *Chest* 1995;108(2 Suppl):43S-52S.
8. Soler N, Torres A, Ewig S, et al. Bronchial microbial patterns in severe exacerbations of chronic obstructive pulmonary disease (COPD) requiring mechanical ventilation. *American journal of respiratory and critical care medicine* 1998;157(5 Pt 1):1498-505. doi: 10.1164/ajrccm.157.5.9711044
9. Monso E, Ruiz J, Rosell A, et al. Bacterial infection in chronic obstructive pulmonary disease. A study of stable and exacerbated outpatients using the protected specimen brush. *American journal of respiratory and critical care medicine* 1995;152(4 Pt 1):1316-20. doi: 10.1164/ajrccm.152.4.7551388
10. Dickson RP, Martinez FJ, Huffnagle GB. The role of the microbiome in exacerbations of chronic lung diseases. *Lancet* 2014;384(9944):691-702. doi: 10.1016/S0140-6736(14)61136-3
11. Wang Z, Bafadhel M, Haldar K, et al. Lung microbiome dynamics in chronic obstructive pulmonary disease exacerbations. *The European respiratory journal* 2016 doi: 10.1183/13993003.01406-2015
12. Molyneaux PL, Mallia P, Cox MJ, et al. Outgrowth of the bacterial airway microbiome after rhinovirus exacerbation of chronic obstructive pulmonary disease. *American journal of respiratory and critical care medicine* 2013;188(10):1224-31. doi: 10.1164/rccm.201302-0341OC
13. Huang YJ, Sethi S, Murphy T, et al. Airway microbiome dynamics in exacerbations of chronic obstructive pulmonary disease. *Journal of clinical microbiology* 2014;52(8):2813-23. doi: 10.1128/JCM.00035-14
14. Millares L, Ferrari R, Gallego M, et al. Bronchial microbiome of severe COPD patients colonised by Pseudomonas aeruginosa. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* 2014;33(7):1101-11. doi: 10.1007/s10096-013-2044-0
15. Huang YJ, Kim E, Cox MJ, et al. A persistent and diverse airway microbiota present during chronic obstructive pulmonary disease exacerbations. *Omics : a journal of integrative biology* 2010;14(1):9-59. doi: 10.1089/omi.2009.0100

16. Huang YJ, Erb-Downward JR, Dickson RP, et al. Understanding the role of the microbiome in chronic obstructive pulmonary disease: principles, challenges, and future directions. *Translational research : the journal of laboratory and clinical medicine* 2017;179:71-83. doi: 10.1016/j.trsl.2016.06.007
17. National Institute for Health and Clinical Excellence. Chronic obstructive pulmonary disease: management of chronic obstructive pulmonary disease in adults in primary and secondary care. London: National Clinical Guideline Centre: Available from: http://guidance.nice.org.uk/CG101/Guidance/pdf/English2010. 2010
18. Caporaso JG, Kuczynski J, Stombaugh J, et al. QIIME allows analysis of high-throughput community sequencing data. *Nature methods* 2010;7(5):335-6. doi: 10.1038/nmeth.f.303
19. Eriksson L JE, Kettaneh-Wold N. Multi- and Megavariate Data Analysis, Part 2, Advanced Applications and Method Extensions. *MKS Umetrics AB* 2006
20. Flores GE, Caporaso JG, Henley JB, et al. Temporal variability is a personalized feature of the human microbiome. *Genome biology* 2014;15(12):531. doi: 10.1186/s13059-014-0531-y
21. R Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.;ISBN 3-900051-07-0, URL http://www.R-project.org/
22. Benjamini Y, and Hochberg, Y. Controlling the false discovery rate – a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol* 1995;57:289-300.
23. Hilty M, Burke C, Pedro H, et al. Disordered microbial communities in asthmatic airways. *PloS one* 2010;5(1):e8578. doi: 10.1371/journal.pone.0008578
24. Erb-Downward JR, Thompson DL, Han MK, et al. Analysis of the lung microbiome in the "healthy" smoker and in COPD. *PloS one* 2011;6(2):e16384. doi: 10.1371/journal.pone.0016384
25. Zakharkina T, Heinzel E, Koczulla RA, et al. Analysis of the airway microbiota of healthy individuals and patients with chronic obstructive pulmonary disease by T-RFLP and clone sequencing. *PloS one* 2013;8(7):e68302. doi: 10.1371/journal.pone.0068302
26. Pragman AA, Kim HB, Reilly CS, et al. The lung microbiome in moderate and severe chronic obstructive pulmonary disease. *PloS one* 2012;7(10):e47305. doi: 10.1371/journal.pone.0047305
27. Dahl M, Vestbo J, Lange P, et al. C-reactive protein as a predictor of prognosis in chronic obstructive pulmonary disease. *American journal of respiratory and critical care medicine* 2007;175(3):250-5. doi: 10.1164/rccm.200605-713OC
28. Langille MG, Zaneveld J, Caporaso JG, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature biotechnology* 2013;31(9):814-21. doi: 10.1038/nbt.2676
29. Worlitzsch D, Tarran R, Ulrich M, et al. Effects of reduced mucus oxygen concentration in airway Pseudomonas infections of cystic fibrosis patients. *The Journal of clinical investigation* 2002;109(3):317-25. doi: 10.1172/JCI13870
30. Schmidt A, Belaaouaj A, Bissinger R, et al. Neutrophil elastase-mediated increase in airway temperature during inflammation. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 2014;13(6):623-31. doi: 10.1016/j.jcf.2014.03.004
31. Marsland BJ, Gollwitzer ES. Host-microorganism interactions in lung diseases. *Nature reviews Immunology* 2014;14(12):827-35. doi: 10.1038/nri3769
32. Lopez-Campos JL, Agusti A. Heterogeneity of chronic obstructive pulmonary disease exacerbations: a two-axes classification proposal. *The Lancet Respiratory medicine* 2015;3(9):729-34. doi: 10.1016/S2213-2600(15)00242-8
33. Wedzicha JA, Brill SE, Allinson JP, et al. Mechanisms and impact of the frequent exacerbator phenotype in chronic obstructive pulmonary disease. *BMC medicine* 2013;11:181. doi: 10.1186/1741-7015-11-181

525 34. Hurst JR, Vestbo J, Anzueto A, et al. Susceptibility to exacerbation in chronic obstructive pulmonary
526 disease. *The New England journal of medicine* 2010;363(12):1128-38. doi:
527 10.1056/NEJMoa0909883

528