Lung microbiome dynamics in chronic obstructive pulmonary disease exacerbations

Zhang Wang[1*], Mona Bafadhel[2*], Koirobi Haldar[3*], Aaron Spivak[1], David Mayhew[1], Bruce E. Miller[4], Ruth Tal-Singer[4], Sebastian L. Johnston[5], Mohammadali Yavari Ramsheh[3], Michael R. Barer[3], Christopher E. Brightling[3€], James R. Brown[1€]

[1] Computational Biology, GSK R&D, Collegeville, Pennsylvania, USA;
[2] Respiratory Medicine Unit, Nuffield Department of Clinical Medicine, University of Oxford, Old Road Campus, Oxford, UK;
[3] Institute for Lung Health, National Institute for Health Research Respiratory Biomedical Research Unit, Department of Infection, Immunity, and Inflammation, University of Leicester, Leicester, UK
[4] Respiratory Therapy Area Unit, GSK R&D, King of Prussia, Pennsylvania, USA;
[5] Airway Disease Infection Section, National Heart and Lung Institute, Imperial College London, London, UK

[*] These authors contributed equally to the manuscript.
[€] Co-senior authors

Correspondence to:
Christopher E. Brightling (ceb17@leicester.ac.uk)
Institute for Lung Health, Clinical Sciences Wing, University Hospitals of Leicester, Leicester, LE3 9QP, UK.
Direct Tel: +441162563340
Secretary Tel: +441162563998

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: 195
Word count text body: 2985

"Take-home" message (119 characters with spaces):
Lung microbiome changes are associated with ~~some~~ COPD exacerbation events and implicated in host inflammatory responses.

Running Head: Lung Microbiome in COPD Exacerbations

Abstract

Increasing evidence suggests that the lung microbiome plays an important role in chronic obstructive pulmonary disease (COPD) severity. However, the dynamics of the lung microbiome during COPD exacerbations and its potential role in disease aetiology remains poorly understood.

We completed a longitudinal 16S ribosomal RNA survey of the lung microbiome on 476 sputum samples collected from 87 subjects with COPD at four visits defined as stable state, exacerbation, two weeks post therapy and six weeks recovery.

Our analysis revealed a dynamic lung microbiota where changes appeared to be associated with exacerbation events and indicative of specific exacerbation phenotypes. Antibiotic and steroid treatments appear to have differential effects on the lung microbiome. We depict a microbial interaction network for the lung microbiome and suggest that perturbation of a few bacterial operational taxonomic units, in particular *Haemophilus* spp, could greatly impact the overall microbial community structure. Furthermore, several serum and sputum biomarkers, in particular sputum interleukin-8 (IL-8), appear to be highly correlated with the structure and diversity of the microbiome.

Our study furthers our understanding of lung microbiome dynamics in COPD patients and highlights its potential as a biomarker, and possibly a target, for future respiratory therapeutics.

Introduction

Chronic obstructive pulmonary disease (COPD), one of the most prevalent respiratory diseases, is characterized by persistent symptoms and impaired lung function as a consequence of airway inflammation, small airway obliteration and alveolar destruction [1-3]. Acute exacerbations of COPD are sudden worsening of symptoms in which bacterial colonization is one major etiologic factor [4-7]. However, the dynamics of the bacterial ecology at exacerbations and its role in disease pathogenesis remains poorly understood.

Advances in next generation sequencing have provided an opportunity to study the lung microbiome in health and disease. Studies using culture-independent techniques such as PCR amplification and sequencing of the 16S ribosomal RNA (rRNA) gene have characterized a distinct bacterial community in the airway of COPD patients compared to healthy subjects and suggest that changes in the lung microbiota could be associated with enhanced airway inflammation and disease progression [8-11]. However, most lung microbiome studies to date involved relatively small cohorts of subjects with limited longitudinal sampling and concurrent clinical information.

We hypothesize that incorporating the lung microbiome profile from larger and better characterized patient cohorts may improve our mechanistic understanding of COPD aetiology as well as provide additional prognostic and therapeutic signatures. Here, we performed a longitudinal 16S rRNA based microbiome survey on 476 sputum samples collected from 87 subjects with COPD from BEAT-COPD (ISRCTN 92422949) [12, 13] which, to our knowledge, is the largest well-characterized COPD lung microbiome cohort to date. We found that changes

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

in the lung microbiome were associated with specific clinical and biochemical characteristics of COPD thereby providing further insights into the relationships among the lung microbial community, host inflammatory responses and disease pathogenesis.

Subjects and methods

Subjects and samples

Sputum samples from COPD subjects were longitudinally collected at four visit types: namely stable state (defined as being eight weeks free from an exacerbation visit), exacerbation (defined according to Anthonisen criteria [14] and/or healthcare utilisation [15]), two weeks post therapy and at recovery (six weeks post exacerbation visit). Exacerbations were treated with oral corticosteroids and antibiotics according to guidelines [16] or study design [13]. Additional details on study subjects, DNA sequencing methodology and biostatical analyses are provided in the online supplement.

Microbiome analysis

Bacterial genomic DNA was extracted from sputum samples using the Qiagen DNA Mini kit (Qiagen, CA, USA) as per manufacture protocol and ~~. T~~the V3-V5 hypervariable regions of the 16S rRNA gene were PCR amplified with the appropriate negative controls. Amplified DNA fragments were ~~and~~ pyrosequenced using ~~ced using~~ 454 Genome Sequencer FLX platform (454 Life Sciences, Roche Diagnostics, UK). Sequencing reads were processed using QIIME pipeline version 1.7 [17]. Stringent criteria were used to remove low quality and chimeric reads. The remaining reads were subject to open reference operational taxonomic unit (OTU) picking (97% identity cutoff). DNA sequencing data is available from the NCBI short reads database, accession No. XXXX. See the Supplemental File for further Materials and Methods content.

Statistical analyses

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

Briefly, exacerbation phenotypes were defined using microbiological and clinical criteria as established previously [12]. Phenotypes of 28 exacerbations samples were undetermined due to missing data. Partial Least Squares Discriminant Analysis (PLS-DA), Receiver Operating Characteristic (ROC) curve reconstruction and network analysis were performed on exacerbation phenotypes and microbiota and/or clinical data. A general linear mixed model (GLMM) was constructed between clinical variables and four measures of alpha diversity (microbial diversity within a sample): OTU richness, Shannon's H, chao1 and Faith's phylogenetic diversity. To identify clinical predictors of beta diversity (microbial composition dissimilarity between samples), canonical correspondence analysis (CCA) was performed on clinical variables and the relative abundance of taxa at the phylum, genus and OTU levels. Biomarker factors were identified using Principal Component Analysis (PCA). All P-values for statistical tests were adjusted for multiple tests using the False Discovery Rate (FDR) [18].

Field Code Changed

Results

Lung microbiome profiles

Sequential sputum samples from 87 subjects were obtained at stable state, exacerbation, two weeks post therapy and six weeks recovery (Fig. 1). At exacerbation, subjects were treated with antibiotics, oral corticosteroids or a combination of both according to guidelines [16] or study design [13]. All samples at exacerbation onset were obtained prior to initiation of treatment. Table 1 shows the clinical characteristics of the subjects at baseline and follow-up visits. Sputum and serum mediator data were collected in stable and exacerbation samples of a proportion of subjects (N = 54) as reported previously (hereafter referred as Group I subjects) [12]. A total of 4,500,748 DNA sequencing reads were generated after demultiplexing and quality control filtering. A rarefaction depth of 1,666 reads per sample was selected based on jackknifed Principal Coordinate Analysis (PCoA) re-sampling analysis (Fig. S1). A total of 1,193 operational taxonomic units (OTUs) were identified across 476 samples.

Field Code Changed

Field Code Changed

Field Code Changed

Based on overall phyla composition, the samples clustered into one of three groups which we refer to as *Proteobacteria*, *Firmicutes* or *Bacteroidetes* subgroups (Fig. S2). About 98.4% sequences belonged to one of four phyla; *Firmicutes* (51.4%), *Proteobacteria* (35.9%), *Actinobacteria* (6.5%) or *Bacteroidetes* (4.6%). Of the 366 genera identified, the most abundant genera were *Streptococcus* (41.1%), *Haemophilus* (18.9%), *Moraxella* (5.6%) and *Pseudomonas* (4.4%), all of which are typical members of the lung microbiota [24]. The genera *Streptococcus* (prevalence: 95.0%), *Haemophilus* (94.7%), *Rothia* (94.1%), *Veillonella* (93.2%) and *Prevotella* (90.3%) were all highly prevalent across all visits (Table 2). There was a significantly greater inter-subject variation in microbiome community at the same visit (weighted UniFrac

0.387±0.186 (s.d.)) compared with temporal variation within each subject (weighted UniFrac 0.272±0.181 (s.d.), $P < 2.2e-16$, T-test) suggesting that for our cohort the lung microbiome was relatively stable over time [8].

Microbiome shifts during exacerbations

Previous reports have emphasized the importance of bacteria in COPD exacerbations [19, 20]~~[18, 19]~~. Our results indicate an overall reduced alpha diversity (microbial diversity within a sample) with a small and non-significant microbial composition shift toward an increase in the relative abundance of *Proteobacteria* (False Discovery Rate or FDR-adjusted for multiple statistical tests ~~([~~adj~~.)~~ .] $P = 0.42$, Paired T-test) and a decrease of *Firmicutes* (adj. $P = 0.73$, Paired T-test) during exacerbations compared to stable samples (Fig. 2a). *Moraxella* showed the greatest change during exacerbations, averaging an increase of relative abundance by 5% (adj. $P = 0.22$, Paired T-test, Fig. 2a,b). This was followed by a decrease of *Streptococcus* (3.8%, adj. $P = 0.58$, Paired T-test) and an increase of *Haemophilus* (3.0%, adj. $P = 0.57$, Paired T-test). Although these changes were not statistically significant in comparisons of paired stable and exacerbation samples, a significant increase of *Moraxella* was observed when comparing exacerbation versus all non-exacerbation samples (adj. $P = 0.022$, T-test).

Further taxonomic breakdown reveals that 95.6% of *Moraxella* DNA sequence reads corresponded to a single OTU (OTU 861881) which has 100% sequence identity to the same 16S rRNA segment from the respiratory pathogen *Moraxella catarrhalis* [21]~~[20]~~. A significant positive correlation was also found between the relative abundance of OTU 861881 and the bacterial load of *M. catarrhalis* as measured by qPCR (Fig. S3). The abundance of *Moraxella*

Field Code Changed

Field Code Changed

Field Code Changed

was most significantly correlated with two host factors, the percentages of sputum neutrophils (adj. P = 0.002, R = 0.18) and sputum macrophages (adj. P = 0.005, R = -0.19).

Despite an overall increase of *Moraxella* during exacerbations across the entire dataset, comparison of paired samples from the same subjects revealed some heterogeneity in *Moraxella* changes (Fig. 2c). Increases in the relative abundance of *Moraxella* during exacerbations was observed in 36 of the 87 subjects due to a potential acquisition of a new *Moraxella* OTU in 23 subjects in whom it was undetectable in the stable samples and a dramatic outgrowth of a pre-existing *Moraxella spp.* by an average of 268 fold (ranging from 2.3-1,412 fold) in another 13 subjects.

Microbiome discriminates between bacterial and eosinophilic exacerbations

Exacerbation phenotypes were defined as either bacterial (number of exacerbation samples N = 33), eosinophilic (N = 19), viral (N = 15), bacterial/eosinophilic combination (N = 3), bacterial/viral combination (N = 12) or pauci-inflammatory (N = 27) using previously published biological criteria [12]. Distinct microbiome profiles at both the phylum and genus levels were observed during exacerbations among subjects across different phenotypes (Fig. 3a). Differences were most pronounced between bacterial and eosinophilic exacerbations, which were more dissimilar from each other in composition than either was from the other subgroups. In particular, there was a significant decrease of alpha diversity (P = 0.04, T-test) and *Firmicutes* (adj. P = 6.3e-5, T-test) and an increase of *Proteobacteria* (adj. P = 2e-4, T-test) in the bacterial subgroup compared to the eosinophilic subgroup (Fig. 3a). At the genus level, this corresponded to a significantly decrease in *Streptococcus* (adj. P = 0.002, T-test) and increase in *Haemophilus* (adj.

Field Code Changed

P = 0.008, T-test) in the bacterial subgroup. Also, a notable decreased *Proteobacteria*:*Firmicutes* ratio was found in eosinophilic subgroups during exacerbations in sharp contrast to all other subgroups (Fig. S4). Furthermore, individual exacerbation samples in the bacterial and eosinophilic subgroups were relatively distinct from each other in both PCoA (Fig. 3b) and unweighted pair group method with arithmetic mean (UPGMA) clustering analyses (Fig. 3c). However, PLS-DA showed only a modest improvement in the prediction of bacterial and eosinophilic exacerbation events by combined clinical and microbiome data versus clinical data alone (Fig. 3d, Table S1).

Oral corticosteroids and antibiotics have different effects on the lung microbiome

Since antibiotics and steroids can alter the lung microbiome in COPD patients [8], we investigated the changes in the microbiome resulting from these treatments. In agreement with an earlier study [8], ~~We~~ we found a decreased microbial alpha diversity with an increase of *Proteobacteria* over *Firmicutes* in subjects treated with corticosteroids alone (Fig. 4). At the genus level, this corresponded to a decrease of *Streptococcus* and an increase of *Haemophilus* and *Moraxella*. An opposite trend in both alpha diversity and microbial composition changes was observed for subjects treated with antibiotics (with or without steroids). The effects of different treatments were further maintained from post therapy to recovery indicating a prolonged effect of treatment on the microbiome.

Network analysis reveals potential microbiota interactions

To gain insight into the interaction between bacterial OTUs in the lung microbiome, we carried out an OTU network analysis using CoNet [22]~~[21]~~. Examination of the microbial network

Field Code Changed

Field Code Changed

revealed that it was predominated by a few "hub" OTUs that were highly connected with multiple other nodes (Fig. 5a, Table S2). For example, the OTU with the highest degree of connectivity in the network was OTU 240755 (*Haemophilus* sp.), which had a co-exclusive relationship with 33 other OTUs. This was followed by OTU 861881 (*Moraxella* sp.), OTU 956702 (*Haemophilus* sp.) and OTU 4445466 (*Streptococcus* sp.), all of which had numerous negative connections with other members of the microbiota. Consequently, abundance increases of OTU 240755, OTU 861881 and OTU 956702 were all associated with a significantly decrease in microbial alpha diversity (Shannon's H vs OTU 240755: adj. P = 1.6e-21, R = -0.43; OTU 861881: adj. P = 8.8e-8, R = -0.24; OTU 956702: adj. P = 1.1e-9, R = -0.28). Other than co-exclusions, co-existence patterns were also observed among bacterial species such as five OTUs showing a strong mutual cooperative relationship in a tightly connected subgraph (Fig. 5a).

Sputum CXCL8/IL-8 as an indicator of community structure and diversity

We ~~next~~ then applied an expanded network analysis which included both bacterial OTUSs and clinical variables to investigate~~d other~~ potential human host interactions with the lung microbiome.~~factors that could potentially mediate microbial interactions by including clinical variables into the network.~~ A set of 66 clinical variables were selected after exclusion of strongly mutually correlated variables. Several sputum mediators, including sputum interleukin-8 (CXCL8/IL-8), MMP-8 and MMP-7 appeared as highly connected nodes in the network (Fig. 5b, Table S3). Among them, sputum CXCL8/IL-8 had the highest degree of microbiota connectivity with a significant negative correlation to 15 OTUs.

We also performed a canonical correspondence analysis (CCA) to test the association of clinical variables with variation in microbial composition at different taxonomic levels. To account for multiple measures per subject, we limited our analysis to the initial samples collected from Group I subjects. Both serum MMP-7 and the percentage of sputum neutrophils had a significant correlation with the microbial composition at the OTU level ($P < 0.05$, CCA, Table 3). Sputum CXCL8/IL-8, the percentage of sputum neutrophils and serum colony stimulating factor 2 (CSF-2) were significantly associated with phylum level variation while sputum CXCL8/IL-8, the percentage of sputum neutrophils and serum MMP-7 were significant at the genus level ($P < 0.05$, CCA, Table 3).

Discussion

COPD occurrence and severity are mediated through complex interactions between the host immune system, environmental factors and respiratory pathogens. Our study adds further insights into the role of the lung microbiome in COPD with the inclusion of a large patient cohort and repeated longitudinal sampling over multiple clinical visits. A key finding is the association of changes in the lung microbiome with multiple characteristics of COPD including specific exacerbation phenotypes, treatment regimen, and the levels of key sputum and serum mediators.

In agreement with previous studies, COPD exacerbation events appear to be associated with decreased microbial diversity and increased proportion of *Proteobacteria* [8, 23]~~[8, 22]~~. In addition, there was a remarkable proliferation of *Moraxella* in a subgroup of subjects during exacerbations (36 out of 87 subjects). *M. catarrhalis* is a critical respiratory pathogen that

Field Code Changed

enhances airway inflammation by stimulating several neutrophil related components during COPD exacerbations [21][~~20~~]. Consistently, there was a significant positive correlation between the abundance of *Moraxella* and the percentage of sputum neutrophils. Despite the heterogeneous nature of COPD, our results suggest that specific subgroups of COPD subjects are particularly susceptible to alternation of microbiome during exacerbations.

Distinct microbial populations existed during exacerbations in subjects with bacterial and eosinophilic exacerbations, which were more dissimilar from each other in microbiome profiles than either was from viral exacerbations. This is in agreement with the suggestion that bacterial and eosinophilic exacerbations reflect fundamental differences in their immunopathogenesis, whereas virus exacerbations are often associated with both bacterial infections and increase of eosinophils [12, 24][~~12, 23~~]. The presence or absence of eosinophilic inflammation could be a potential biomarker for stratification of the underlying associated microbiome.

Our study shows that current standard of care treatments potentially alter the lung microbiome. In particular, a reduction in microbial diversity and an increased *Proteobacteria:Firmicutes* ratio toward recovery were observed in subjects treated with steroids alone, whereas the trend was reversed in subjects with antibiotics. Similar changes in the lung microbiome were observed in a smaller study of twelve COPD subjects after these treatments [8]. In support of reports on the limited efficacy and greater side effects of steroids [25, 26][~~24, 25~~] our results suggest that steroids alone could affect the lung microbiome and underscore the importance of patient stratification approaches, such as blood eosinophil guided prednisolone therapy [13] toward more precise drug management strategies.

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

~~Widespread and extensive interactions between individual bacterial species are evident in other body sites [21].~~ In our network analysis, several bacterial OTUs, in particular the OTU 240755 belonging to *Haemophilus*, were identified as microbial "hubs" that had a disproportionately large number of negative connections with other OTUs. Such correlations were highly robust across all samples, indicating this might represent a general pattern of the COPD lung microbiome. Overgrowth of these bacterial OTUs could thus drive respiratory tract dysbiosis which has been suggested to be a potential cause of lung disease exacerbations [27]~~[26]~~. Recent studies have highlighted the importance of ecological interactions in multiple human body habitats. An emerging paradigm is the "keystone species" hypothesis where even marginal changes in the abundances of relatively few bacterial species could have profound effects on the overall microbial community structure and, consequently, alter human disease states [22, 28, 29]~~[21, 27, 28]~~. Based on our network analyses we speculate that increased abundance of *Haemophilus* sp, and possibly other proteobacteria, might remodel the normal lung microbial ecosystem into a state of dysbiosis which could elicit a host pro- inflammatory response. Our results suggest that the "keystone species" hypothesis in context of the lung microbiome warrants further study since it might provide a conceptual basis for novel therapeutic strategies that target a few key bacterial targets to counteract a dysbiotic microbial community in COPD.

Sputum CXCL8/IL-8 was significantly associated with both lung microbiome diversity and its overall community structure, thus CXCL8/IL-8 could be a suitable biomarker to monitor the overall lung microbial population. Sputum CXCL8/IL-8 has long been recognized to play a prominent role in COPD [30, 31]~~[29, 30]~~. It induces airway inflammation by predominantly recruiting neutrophils and upregulating airway mucin genes resulting in mucus production

Field Code Changed

Field Code Changed

Formatted: Font: Italic

Field Code Changed

[32][~~31~~]. Elevated sputum CXCL8/IL-8 levels are associated with elevated COPD severity [30, 33, 34][~~29, 32, 33~~]. Pathogenic members of *Haemophilus* and *Moraxella* are able to directly induce inflammation by exposing the host to lipopolysaccharides and other pathogen-associated molecular patterns (PAMPs) [35][~~34~~]. Interestingly, our network analysis revealed that multiple members in the lung microbiome were negatively correlated with both sputum CXCL8/IL-8 including several OTUs comprised of *Haemophilus*, *Moraxella* and *Streptococcus* species. Potentially, these pathogens could also indirectly trigger the excessive production of CXCL8/IL-8 through dysbiosis of the lung microbiota. Therefore, we speculate that the lung microbiome could serve as an additional line of defence that shapes the lung inflammatory response induced by respiratory pathogens [27][~~26~~].

There are several caveats to our study. First, despite a large cohort size, our survey focused exclusively on COPD patients with exacerbations with no healthy or non-exacerbator control subjects. Data from these populations could be informative in defining the normal lung microbiota as well as the changes of microbial composition resulting in COPD onset. Second, besides bacteria, the importance of viruses and fungi in COPD is just beginning to be appreciated and studied [36-38][~~35-37~~] . A systems biology view of bacterial, viral and fungal microbiomes integrating additional host response factors such as host transcriptome and metabolome profiles would boost our understanding of the host-microbiota interaction and its implication in disease aetiology. Finally, our results need to be replicated in further larger and distinct patient populations including those with different ethnicity and biogeographical backgrounds.

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

In conclusion, we show that changes in the lung microbiome are associated with COPD exacerbation events and are potentially implicated in mediating host inflammatory responses in some subjects. Moreover, this work furthers our understanding of the lung microbiome in COPD and opens potential avenues for new biomarkers and respiratory therapeutics.

Acknowledgement

The study was sponsored by University Hospitals of Leicester NHS Trust. Study cohort BEAT-COPD was co-funded by Medical Research Council, UK and AstraZeneca with microbiome analysis funded by GSK. CEB and MB were funded by the National Institute for Health Research. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. ZW is supported by the GSK Early Talent Postdoctoral Fellowship Program. We thank Deepali Jhamb, and Michal Magid-Slav of Computational Biology, GSK for their advice on statistic analyses of this study.

References

1. Lopez AD, Shibuya K, Rao C, *et al*. Chronic obstructive pulmonary disease: current burden and future projections. *Eur Respir J* 2006 Feb; 27: 397-412.

2. FRANKLIN W, LOWELL FC, MICHELSON AL, SCHILLER IW. Chronic obstructive pulmonary emphysema; a disease of smokers. *Ann Intern Med* 1956 Aug; 45: 268-274.

3. Taraseviciene-Stewart L, Douglas IS, Nana-Sinkam PS, *et al*. Is alveolar destruction and emphysema in chronic obstructive pulmonary disease an immune disease? *Proc Am Thorac Soc* 2006 Nov; 3: 687-690.

4. Miravitlles M, Espinosa C, Fernandez-Laso E, Martos JA, Maldonado JA, Gallego M. 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 Jul; 116: 40-46.

5. Ball P. Epidemiology and treatment of chronic bronchitis and its exacerbations. *Chest* 1995 Aug; 108: 43S-52S.

6. Soler N, Torres A, Ewig S, *et al*. Bronchial microbial patterns in severe exacerbations of chronic obstructive pulmonary disease (COPD) requiring mechanical ventilation. *Am J Respir Crit Care Med* 1998 May; 157: 1498-1505.

7. 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. *Am J Respir Crit Care Med* 1995 Oct; 152: 1316-1320.

8. Huang YJ, Sethi S, Murphy T, Nariya S, Boushey HA, Lynch SV. Airway microbiome dynamics in exacerbations of chronic obstructive pulmonary disease. *J Clin Microbiol* 2014 Aug; 52: 2813-2823.

9. Molyneaux PL, Mallia P, Cox MJ, *et al*. Outgrowth of the Bacterial Airway Microbiome following Rhinovirus Exacerbation of Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med* 2013 Aug 30.

10. 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: e68302.

11. Rogers GB, Carroll MP, Serisier DJ, Hockey PM, Jones G, Bruce KD. characterization of bacterial community diversity in cystic fibrosis lung infections by use of 16s ribosomal DNA terminal restriction fragment length polymorphism profiling. *J Clin Microbiol* 2004 Nov; 42: 5176-5183.

12. Bafadhel M, McKenna S, Terry S, *et al*. Acute exacerbations of chronic obstructive pulmonary disease: identification of biologic clusters and their biomarkers. *Am J Respir Crit Care Med* 2011 Sep 15; 184: 662-671.

13 Bafadhel M, McKenna S, Terry S, *et al*. Blood eosinophils to direct corticosteroid treatment of exacerbations of chronic obstructive pulmonary disease: a randomized placebo-controlled trial. *Am J Respir Crit Care Med* 2012 Jul 1; 186: 48-55.

14 Anthonisen NR, Manfreda J, Warren CP, Hershfield ES, Harding GK, Nelson NA. Antibiotic therapy in exacerbations of chronic obstructive pulmonary disease. *Ann Intern Med* 1987 Feb; 106: 196-204.

15 Rodriguez-Roisin R. Toward a consensus definition for COPD exacerbations. *Chest* 2000 May; 117: 398S-401S.

16 National Institute for Health and Clinical Excellence. Chronic obstructive pulmonary disease: management of chronic obstructive pulmonary disease in adults in primary and secondary care. 2010. London, National Institute for Health and Clinical Excellence.
Ref Type: Report

17 Caporaso JG, Kuczynski J, Stombaugh J, *et al*. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010 May; 7: 335-336.

18 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach for multiple testing. *Journal of the Royal Statistical Society, Series B* 1995 Jan 1; 57: 289-300.

19 Erkan L, Uzun O, Findik S, Katar D, Sanic A, Atici AG. Role of bacteria in acute exacerbations of chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis* 2008; 3: 463-467.

20 Sethi S, Murphy TF. Bacterial infection in chronic obstructive pulmonary disease in 2000: a state-of-the-art review. *Clin Microbiol Rev* 2001 Apr; 14: 336-363.

21 Parameswaran GI, Wrona CT, Murphy TF, Sethi S. Moraxella catarrhalis acquisition, airway inflammation and protease-antiprotease balance in chronic obstructive pulmonary disease. *BMC Infect Dis* 2009; 9: 178.

22 Faust K, Sathirapongsasuti JF, Izard J, *et al*. Microbial co-occurrence relationships in the human microbiome. *PLoS Comput Biol* 2012; 8: e1002606.

23 Millares L, Ferrari R, Gallego M, *et al*. Bronchial microbiome of severe COPD patients colonised by Pseudomonas aeruginosa. *Eur J Clin Microbiol Infect Dis* 2014 Jan 22.

24 Bathoorn E, Kerstjens H, Postma D, Timens W, MacNee W. Airways inflammation and treatment during acute exacerbations of COPD. *Int J Chron Obstruct Pulmon Dis* 2008; 3: 217-229.

25 Leuppi JD, Schuetz P, Bingisser R, *et al*. Short-term vs conventional glucocorticoid therapy in acute exacerbations of chronic obstructive pulmonary disease: the REDUCE randomized clinical trial. *JAMA* 2013 Jun 5; 309: 2223-2231.

26 McEvoy CE, Niewoehner DE. Adverse effects of corticosteroid therapy for COPD. A critical review. *Chest* 1997 Mar; 111: 732-743.

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Space After: 12 pt

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

27 Dickson RP, Martinez FJ, Huffnagle GB. The role of the microbiome in exacerbations of chronic lung diseases. *Lancet* 2014 Aug 23; 384: 691-702.

28 Fisher CK, Mehta P. Identifying keystone species in the human gut microbiome from metagenomic timeseries using sparse linear regression. *PLoS One* 2014; 9: e102451.

29 Hajishengallis G, Darveau RP, Curtis MA. The keystone-pathogen hypothesis. *Nat Rev Microbiol* 2012 Oct; 10: 717-725.

30 Larsson K. Inflammatory markers in COPD. *Clin Respir J* 2008 Oct; 2 Suppl 1: 84-87.

31 Chung KF. Cytokines in chronic obstructive pulmonary disease. *Eur Respir J Suppl* 2001 Dec; 34: 50s-59s.

32 Bautista MV, Chen Y, Ivanova VS, Rahimi MK, Watson AM, Rose MC. IL-8 regulates mucin gene expression at the posttranscriptional level in lung epithelial cells. *J Immunol* 2009 Aug 1; 183: 2159-2166.

33 Zhang X, Zheng H, Zhang H, *et al*. Increased interleukin (IL)-8 and decreased IL-17 production in chronic obstructive pulmonary disease (COPD) provoked by cigarette smoke. *Cytokine* 2011 Dec; 56: 717-725.

34 Bhowmik A, Seemungal TA, Sapsford RJ, Wedzicha JA. Relation of sputum inflammatory markers to symptoms and lung function changes in COPD exacerbations. *Thorax* 2000 Feb; 55: 114-120.

35 Ishii KJ, Koyama S, Nakagawa A, Coban C, Akira S. Host innate immune receptors and beyond: making sense of microbial infections. *Cell Host Microbe* 2008 Jun 12; 3: 352-363.

36 Kistler A, Avila PC, Rouskin S, *et al*. Pan-viral screening of respiratory tract infections in adults with and without asthma reveals unexpected human coronavirus and human rhinovirus diversity. *J Infect Dis* 2007 Sep 15; 196: 817-825.

37 Cui L, Lucht L, Tipton L, *et al*. Topographic diversity of the respiratory tract mycobiome and alteration in HIV and lung disease. *Am J Respir Crit Care Med* 2015 Apr 15; 191: 932-942.

38 Ghannoum MA, Jurevic RJ, Mukherjee PK, *et al*. Characterization of the oral fungal microbiome (mycobiome) in healthy individuals. *PLoS Pathog* 2010 Jan; 6: e1000713.

~~References~~

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Font: Italic, Do not check spelling or grammar

Formatted: Indent: Left: 0 cm, Hanging: 1.27 cm, Line spacing: single, Tab stops: 0.95 cm, Right + 1.27 cm, Left

Figure Legends

Figure 1. Flow diagram for BEAT-COPD subject enrolment.

Figure 2. Microbiome shifts during exacerbations. (**a**) Alpha diversity (Faith's phylogenetic diversity) and composition of major taxonomic groups at both phylum and genus levels in samples collected across

the four visit types stable, exacerbation (Exac), post therapy (Post) and recovery (Rec). The number of samples is indicated for each subgroup in the bar chart. (**b**) Box and Whisker plots showing the relative abundances of *Streptococcus*, *Haemophilus* and *Moraxella* in samples collected across the four visits. (**c**). Heterogeneity in *Moraxella* changes among individuals. Lines connect paired stable and exacerbation samples collected in the same visit series from the same subjects, and were colored by increase (red) or decrease (green) of *Moraxella* during exacerbations. Only paired stable and exacerbation samples were included.

Figure 3. The microbiome discriminates bacterial and eosinophilic exacerbations. (**a**) Alpha diversity (Shannon's H) and composition of major taxonomic groups at both phylum and genus levels in exacerbation samples with different exacerbation phenotypes. The number of samples is indicated for each subgroup in the bar chart. Alpha diversity or genus level taxonomic groups that differed significantly between bacterial and eosinophilic exacerbations were indicated in asterisks ($P < 0.05$, T-test). (**b,c**) PCoA (**b**) and UPGMA clustering (**c**) show distinct clustering of samples in bacterial and eosinophilic subgroups. (**d**) PLS-DA classification of bacterial and eosinophilic exacerbations using clinical, microbiome and their combined variables at both phylum (L6) and OTU levels. The models were evaluated in terms of AUC, R^2 and Q^2 scores. Letters indicate exacerbation classification as: bacterial (B); viral (V); eosinophilic (E); bacterial and eosinophilic (BE); bacterial and viral (BV) and; pauci-inflammatory (Pauci).

Figure 4. Microbiome changes due to oral corticosteroids and antibiotics. Alpha diversity (Shannon's H) and composition of major taxonomic groups at both phylum and genus levels in exacerbations (Exac), post therapy (Post) and recovery (Rec) samples of subjects treated by steroids, antibiotics or a combination of both. Alpha diversity or genus level taxonomic groups that differed significantly subsequent to particular treatments were indicated in asterisks ($P < 0.05$, ANOVA). Only visit series with a complete cycle of exacerbation, post therapy and recovery visits were included. The number of visit

series is indicated for each subgroup.

Figure 5. Bacterial co-existence and co-exclusion relationships with OTUs and host factors. Interaction networks of (**a**) microbiome and (**b**) microbiome and clinical factors. Each node represents an OTU or a clinical trait. The OTUs were colored by the class-level taxonomy. The five co-existence OTUs are highlighted in the dotted eclipse in (**a**). The clinical traits were grouped together at the left of the network in (**b**). Each edge represents a significant correlation coloured by co-existence (green) or co-exclusion relationships (red). Edge width is proportional to the absolute value of Pearson correlation coefficient. The size of the node is proportional to its degree of connectivity. The degrees are shown in parentheses for highly connected nodes. For the clinical variables, both the total degree and the degree of connectivity to OTUs are shown.

Table 1. Major clinical characteristics of subjects at baseline and over four visits.

	All subjects		Visits				P-Value
			Stable (N=106)	Exacerbation (N=137)	Post therapy (N=136)	Recovery (N=97)	
Sex †	Male (65), Female (22)	FEV1 ‡	1.3 (0.1)	1.1 (0.1)	1.2 (0.1)	1.2 (0.1)	0.1
Age [1]	68 (45-87)	FEV1% predicted	49.6 (1.8)	48.1 (1.6)	48.5 (1.6)	49.0 (2.0)	0.94
Age at diagnosis [1]	61 (30-83)	FEV1/FVC ratio	48.0 (1.3)	48.9 (1.3)	49.2 (1.2)	48.7 (1.4)	0.92
BMI [1]	26.40 (16.67-38.19)	CRQ total	16.6 (0.5)	12.7 (0.4)	16.0 (0.4)	16.6 (0.5)	< 0.001
Baseline § GOLD status	1 (1), 2 (35), 3 (32), 4 (19)	VAS total	174.2 (8.5)	255.7 (6.5)	150.2 (7.5)	151.7 (8.5)	< 0.001
Smoking status	current smoker (37), ex smoker (48), non smoker (2)	Sputum pathogen detection	47	55	42	32	0.30
Pack year history [1]	50 (6-158)	Sputum cell count ($X10^6$ cells/g sputum) [2]	5.8 (4.4-7.3)	14.6 (11.7-17.5)	8.1 (5.7-10.4)	5.1 (3.6-6.6)	< 0.001 *
Number of exacerbations	1 (46), 2 (31), 3 (9), 4 (1)	Sputum neutrophil count % [2]	69.1 (64.8-73.5)	79.1 (75.4-82.9)	74.2 (70.6-77.8)	68.7 (63.9-73.5)	< 0.001
Treatment ‖	antibiotics (21), steroid (8), both (65) ††	Sputum eosinophil count % [2]	2.8 (1.9-3.7)	4.3 (2.6-6.0)	1.5 (1.0-2.0)	3.8 (2.3-5.3)	< 0.01 *
Baseline FEV1	1.3 (0.1)	Sputum lymphocyte count % [2]	0.5 (0.4-0.7)	0.5 (0.4-0.6)	0.6 (0.4-0.7)	0.5 (0.4-0.7)	0.91
Baseline FEV1%	47.4 (2.0)	Blood cell count ($X10^9$	8.3 (7.7-8.9)	9.5 (8.8-10.2)	10.4 (9.5-11.2)	8.8 (8.1-9.6)	< 0.01

predicted		cells/L) [2]					
Baseline FEV1/FVC ratio	46.7 (1.4)	Blood neutrophil ($X10^9$ cells/L) [2]	5.5 (5.1-5.9)	6.9 (6.4-7.4)	7.6 (7.1-8.2)	5.9 (5.4-6.4)	< 0.001
Baseline CRQ [€] total, units	16.2 (0.5)	Blood eosinophil ($X10^9$ cells/L) [2]	0.2 (0.2-0.3)	0.3 (0.2-0.3)	0.2 (0.2-0.3)	0.3 (0.2-0.3)	0.16
Baseline VAS [€] total, mm	159.6 (8.5)	Blood lymphocyte ($X10^9$ cells/L) [2]	2.1 (1.9-2.3)	2.1 (2.0-2.3)	2.3 (2.1-2.5)	2.2 (2.0-2.5)	0.29
Baseline SGRQ [€] total, units	52.9 (1.9)	Blood basophil ($X10^9$ cells/L) [2]	0.04 (0.03-0.04)	0.04 (0.04-0.04)	0.04 (0.04-0.04)	0.04 (0.04-0.05)	0.62

[†] Categorical data present as category (number).

[‡] Continuous data present as mean (SEM) unless stated below.

[1] Mean (range). [2] Mean (95% confidence interval).

[§] Baseline visits prior to the stable visits for sputum collection.

[ll] Treatments administered for exacerbations. Assessments at exacerbation were prior to initiation of therapy.

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

[€] CRQ = Chronic Respiratory Disease Questionnaire; VAS = Visual analog score; SGRQ = St. George's Respiratory Questionnaire.

* These variables were log transformed for statistic analysis.

Table 2. The prevalence (P) and average relative abundance (RA) of predominant OTUs (average relative abundance > 1%) in the lung microbiome. The OTUs were firstly grouped by their genera and then ranked by their RAs.

OTU ID †	**Species/Subspecies**	**All (N=476)**		**Stable (N=106)**		**Exacerbation (N=137)**		**Post therapy (N=136)**		**Recovery (N=97)**	
		P	**RA**	**P**	**RA**	**P**	**RA**	**P**	**RA**	**P**	**RA**
4439603	*Streptococcus sp.*	82.1	13.1	83.0	12.3	78.8	11.6	82.4	13.4	85.6	15.8
4445466	*Streptococcus sp.*	44.7	10.1	47.2	12.1	45.3	8.9	46.3	10.1	39.2	9.5
509773	*Streptococcus sp.*	78.4	6.6	85.8	7.3	80.3	6.3	71.3	6.2	77.3	6.9
1059655	*Streptococcus sp.*	31.5	5.2	30.2	5.2	33.6	5.4	30.9	4.9	30.9	5.6
4462083	*Streptococcus infantis*	31.7	3.9	29.2	4.3	32.1	4.0	31.6	4.2	34.0	3.0
240755	*Haemophilus sp.*	69.7	10.7	61.3	9.5	72.3	12.4	71.3	9.5	73.2	11.3
956702	*Haemophilus sp.*	25.2	5.1	33.0	6.5	24.1	5.1	23.5	5.4	20.6	3.0
4385138	*Haemophilus sp.*	23.5	1.4	25.5	0.8	27.0	1.9	22.1	0.9	18.6	1.9
861881	*Moraxella sp.*	46.2	5.6	45.3	5.0	57.7	10.0	41.9	3.4	37.1	3.3
269930	*Pseudomonas veronii*	33.8	2.4	42.5	3.0	32.8	2.6	32.4	2.2	27.8	1.8
269901	*Pseudomonas sp.*	12.8	1.6	14.2	1.8	12.4	1.2	14.7	1.4	9.3	2.4
342427	*Veillonella dispar*	93.3	2.7	91.5	2.4	93.4	2.2	94.9	3.2	92.8	3.0
4326277	Unclassified in *Gemellaceae*	84.0	2.6	88.7	2.8	81.8	2.3	78.7	2.4	89.7	3.0
4411138	*Rothia mucilaginosa*	73.1	2.4	75.5	3.0	74.5	2.3	71.3	2.1	71.1	2.5
4396235	*Neisseria sp.*	58.6	2.2	60.4	2.3	59.9	2.2	55.9	2.4	58.8	1.8
12574	*Actinomyces sp.*	79.2	1.6	84.0	1.7	79.6	1.3	72.8	1.7	82.5	1.7
257492	*Granulicatella sp.*	83.8	1.4	86.8	1.4	85.4	1.2	80.9	1.3	82.5	1.7
4307391	*Prevotella melaninogenica*	39.1	1.1	41.5	1.1	39.4	1.2	40.4	1.2	34.0	0.8

Table 3. List of clinical variables significantly associated with microbial alpha and beta diversity in Group I subjects.

Alpha diversity	**Shannon's H**	**Observed Species**	**Faith's Phylogenetic Diversity**	**Chao1**	**Positive/Negative correlation**
Sputum CXCL8/IL-8	***	*	***	***	N
Baseline SGRQ † symptom	-	***	***	***	N
BMI	*	*	*	-	P
Sputum CXCL11/ITAC	-	*	*	**	P
Serum TNF-α	*	*	-	-	P
Serum SAA-1	-	*	*	-	P
Serum CCL26/eotaxin3	*	-	-	-	N
Serum CSF-2	*	-	-	-	P
Serum IL-10	*	-	-	-	P
Sputum MMP-8	-	*	-	-	N
Blood monocytes	-	-	**	-	P
Blood basophils	-	-	*	-	P

Beta diversity	**OTU**	**Genus (L6)**	**Phylum (L2)**
Sputum CXCL8/IL-8	-	**	**
Serum MMP-7	**	**	-
Sputum neutrophil percentage	**	**	**
Serum CSF-2	-	-	*
BMI	-	-	- ‡
Exacerbation frequency	-	-	- ‡

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; - $p \geq 0.05$ and absent in the model.

† SGRQ = St. George's Respiratory Questionnaire.

‡ BMI and Exacerbation frequency are not significantly associated with L2 beta diversity but present in its model.